



HAL
open science

Quantifier le carbone organique du sol par spectroscopie infrarouge

Bernard G Barthès, Aurélie Cambou, Tiphaine Chevallier

► **To cite this version:**

Bernard G Barthès, Aurélie Cambou, Tiphaine Chevallier. Quantifier le carbone organique du sol par spectroscopie infrarouge. *Étude et Gestion des Sols*, 2023, 30, pp.287-305. hal-04515788

HAL Id: hal-04515788

<https://hal.science/hal-04515788v1>

Submitted on 21 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Quantifier le carbone organique du sol par spectroscopie infrarouge

Bernard G. Barthès^{1,2}, Aurélie Cambou^{1,2*}, Tiphaine Chevallier^{1,2}

1. IRD, UMR Eco&Sols, 2 place Viala, 34060 Montpellier Cedex 2, France

2. Eco&Sols, Université de Montpellier, Cirad, Inrae, IRD, Institut Agro, 34060 Montpellier, France

*Auteure correspondante : Aurélie Cambou (aurelie.cambou@ird.fr)

Titre courant : quantification infrarouge de C du sol

Résumé

Le carbone organique du sol (COS) est un pilier des services écosystémiques rendus par les sols. Cependant, quantifier le COS, notamment en termes de stock ($\text{kgC}\cdot\text{m}^{-2}$ pour une couche de sol), ou caractériser les fractions qui le composent, pour mieux comprendre sa dynamique, sont souvent fastidieux et/ou coûteux. Cet article propose une vue d'ensemble présentant l'intérêt de la spectroscopie infrarouge (IR) pour quantifier à haut débit et faible coût le COS et ses fractions.

De nombreux travaux montrent que la spectroscopie IR permet de quantifier précisément la concentration en COS ($\text{gC}\cdot\text{kg}^{-1}$). Dans cette approche, un modèle statistique est construit sur un jeu d'échantillons d'étalonnage pour exprimer la concentration en COS (dosée de manière conventionnelle) en fonction du spectre ; puis ce modèle est appliqué au spectre de nouveaux échantillons pour prédire leur concentration en COS. En général le domaine moyen IR (MIR, 2500-25 000 nm) permet de meilleures prédictions pour les sols des régions tempérées lorsque les spectres sont acquis sur matériau finement broyé ($< 0,2$ mm), tandis que le domaine proche IR (PIR, 800-2500 nm) permet de meilleures prédictions pour les sols des régions chaudes ou lorsque les spectres sont acquis sur matériau plus grossièrement préparé (< 2 mm) ou sur le terrain. En effet, les spectres peuvent être acquis en conditions de laboratoire (sur sol sec < 2 ou $< 0,2$ mm), mais également sur le terrain (sur carottes, sol émiétté, *etc.*). La spectroscopie de terrain élimine les étapes de conditionnement, transport et préparation des échantillons (sauf pour l'étalonnage), mais produit des prédictions moins

précises en général, du fait notamment de l'humidité variable entre échantillons. De plus, la construction d'étalonnages avec des spectres de terrain requiert jusqu'à présent des campagnes spécifiques de prélèvement et d'analyse de sol ; elle est donc beaucoup plus lourde qu'avec des spectres de laboratoire, car des collections d'échantillons secs < 2 mm déjà analysés sont souvent accessibles, sur lesquels il suffit d'acquérir des spectres pour construire des étalonnages.

Des travaux récents montrent que la spectroscopie IR permet également de prédire le stock de COS, sans avoir à déterminer la densité apparente (sauf pour les étalonnages). Les prédictions sont souvent plus précises à partir de spectres de sol sec < 2 mm qu'à partir de spectres de terrain, sauf quand la teneur en éléments grossiers (> 2 mm) est élevée et varie entre échantillons. Enfin, la distribution du COS entre ses différentes fractions, par exemple granulométriques, peut être prédite à partir du spectre IR du sol total. La spectroscopie IR permet ainsi de faciliter la caractérisation du COS, et de valoriser les analyses conventionnelles réalisées sur certains échantillons en les utilisant pour des prédictions sur d'autres échantillons.

Mots-clés

Matière organique du sol, stock de carbone du sol, proche infrarouge, moyen infrarouge, régressions multivariées

Quantifying soil organic carbon by infrared spectroscopy

Abstract

Soil organic carbon (SOC) is a pillar of soil ecosystem services. However, SOC quantification, especially in terms of stock (in kgC m^{-2} for a given soil layer), or SOC fraction characterization, which helps understand SOC dynamics, are often tedious and/or costly. This article proposes an overview to show the value of IR spectroscopy to quantify SOC and its fractions time- and cost-effectively.

Many works have demonstrated that IR spectroscopy allows accurate determination of SOC concentration (gC kg^{-1}). In this approach, a statistical model is built on a set of calibration samples in order to express SOC (determined conventionally) as a function of the spectrum; then the model is applied on the spectra of new samples to predict their SOC concentration. In general, the mid-IR domain (MIR, 2,500-25,000 nm) yields better predictions for soils from

temperate regions when spectra are acquired on finely ground samples (< 0.2 mm), while the near IR domain (NIR, 800-2500 nm) yields better predictions for soils from (sub)tropical regions or when spectra are acquired on more coarsely prepared samples (< 2 mm) or in the field. Indeed, spectrum acquisition is possible in the laboratory (on dry soil < 2 or < 0.2 mm) but also in the field (on cores, clods, etc.). Field spectroscopy eliminates the steps of sample packaging, transport and preparation (except for calibration), but yields less accurate predictions in general, mainly due to variable moisture between samples. Moreover, calibration with field spectra requires specific campaigns of soil sampling and analysis, so is much heavier than calibration with laboratory spectra; because collections of dry samples < 2 mm already analyzed conventionally are often available, so that building calibration models only requires spectrum acquisition on these samples.

Recent works have also shown the interest of IR spectroscopy for predicting SOC stock, without having to determine bulk density (except for calibration). Predictions are more accurate in general when using spectra of dry soil < 2 mm than field spectra, except when coarse particle content (> 2 mm) is high and varies between samples. Furthermore, the IR spectrum of bulk soil can be used for predicting SOC distribution among its fractions, for instance particle size fractions. So IR spectroscopy facilitates SOC characterization, and increases the value of conventional analyses carried out on some samples by using them for predictions on other samples.

Key-words

Soil organic matter, soil carbon stock, near infrared, mid infrared, multivariate regressions

Cuantificación del carbono orgánico del suelo mediante espectroscopia infrarroja

Resumen

El carbono orgánico del suelo (COS) es un pilar de los servicios ecosistémicos del suelo; pero su cuantificación es a menudo tediosa y/o costosa, especialmente en términos de stock de COS (en $\text{kgC}^\circ\text{m}^{-2}$ para una capa de suelo), como lo es también la caracterización de las fracciones que componen el COS. Este artículo propone una visión general para mostrar el valor de la espectroscopia infrarroja (IR) para cuantificar el COS y sus fracciones de una forma rápida y rentable.

Muchos trabajos han demostrado que la espectroscopia IR permite determinar con precisión la concentración de COS (gC kg^{-1}). En este enfoque, se construye un modelo estadístico sobre

un conjunto de muestras de calibración con el fin de expresar el COS en función del espectro; a continuación, el modelo se aplica sobre los espectros de nuevas muestras para predecir su concentración de COS. En general, el dominio del IR medio (MIR, 2500-25 000 nm) arroja mejores predicciones para los suelos de regiones templadas cuando las muestras son finamente molidas (< 0,2 mm), mientras que el dominio del IR (visible y) cercano (CIR, 800-2500 nm; VCIR, 350-2500 nm) arroja mejores predicciones para los suelos de regiones (sub)tropicales o más toscamente preparados (< 2 mm), pero también en el campo. En efecto, la adquisición del espectro es posible en el laboratorio (en suelos secos < 2 o < 0,2 mm) pero también en el campo (en paredes de fosas, núcleos, terrones, etc.). La espectroscopia de campo elimina las etapas de embalaje, transporte y preparación de las muestras (excepto para la calibración), pero produce predicciones menos precisas en general, principalmente debido a la humedad variable entre las muestras. Además, la calibración con espectros de campo requiere campañas específicas de muestreo y análisis del suelo, por lo que es mucho más pesada que la calibración con espectros de laboratorio ; porque a menudo se dispone de colecciones de muestras secas < 2 mm ya analizadas convencionalmente, de modo que la construcción de calibraciones sólo requiere la adquisición de espectros en estas muestras.

Trabajos recientes han mostrado también el interés de la espectroscopia IR para predecir el stock de COS, sin tener que determinar la densidad aparente (excepto para la calibración). En general, las predicciones son más precisas cuando se utilizan espectros de suelo seco < 2 mm que los espectros de campo, excepto cuando el contenido de partículas gruesas (> 2 mm) es elevado y varía entre las muestras. Por último, la distribución del COS entre sus diferentes fracciones, por ejemplo la granulométrica, puede predecirse a partir del espectro IR del suelo total. Así pues, la espectroscopia IR facilita la caracterización del COS y aumenta el valor de los análisis convencionales realizados en algunas muestras al utilizarlos para hacer predicciones en otras.

Palabras clave

Materia orgánica del suelo, stock de carbono del suelo, infrarrojo cercano, infrarrojo medio, regresiones multivariantes

1. Introduction

Le carbone organique, élément majeur de la matière organique des sols, est l'un des principaux piliers des services écosystémiques rendus par les sols. En effet, il joue un rôle prépondérant et positif sur leur fertilité physique (agrégation, rétention en eau, aération, etc.), chimique (fourniture de nutriments, pouvoir tampon, etc.) et biologique (activité et diversité microbienne, animale et végétale ; Lal, 2014). De plus, les phénomènes de stockage/déstockage du carbone organique du sol (COS) ont un effet direct sur la concentration en gaz carbonique (CO₂) dans l'atmosphère, et donc sur le climat (GIEC, 2019). Le temps de stockage du COS est très variable, allant de la journée au millier d'années, en fonction de sa nature, de son environnement (physique, chimique et biologique) mais aussi des perturbations anthropiques, auxquelles il est particulièrement sensible, le rendant, de ce fait, gérable (Derrien *et al.*, 2016). C'est dans ce contexte que des initiatives ont été lancées (par exemple l'Initiative Internationale « 4 pour 1000 », 2018) pour promouvoir des pratiques agricoles dites « stockantes », afin de maintenir voire augmenter la quantité de COS. Cette dernière doit donc pouvoir être mesurée, ce qui requiert des méthodes de quantification.

La quantification conventionnelle de la concentration en COS est le plus souvent réalisée par combustion sèche, parfois encore par oxydation sulfochromique, qui toutes deux requièrent des réactifs et produisent des déchets relativement toxiques (Pansu et Gautheyrou, 2006). Toutefois, la concentration en COS (en gC.kg⁻¹) n'est pas une variable additive (la concentration en COS de la somme de deux éléments, horizons ou parcelles par exemple, n'est pas la somme des concentrations en COS de chaque élément). Aussi, quantifier le carbone du sol à l'échelle d'un territoire requiert de déterminer la masse volumique apparente (ou densité apparente, Da, en kg.dm⁻³), afin de calculer le stock de COS (en gC.dm⁻³, ou en kgC.m⁻² pour une couche de sol donnée), qui est une variable additive (Poeplau *et al.*, 2017 ; dans le calcul du stock, le produit de la concentration, déterminée sur la terre fine < 2 mm, et de la Da, déterminée sur le sol total, est ensuite pondéré par la proportion massique de terre fine dans le sol total, relativement simple à mesurer). La densité apparente est généralement déterminée en pesant après séchage à 105°C un volume connu de sol intact (à structure conservée), le plus souvent prélevé à l'aide d'un cylindre biseauté, ce qui nécessite l'ouverture d'une fosse ou l'utilisation d'un dispositif de

carottage mécanisé. Quantifier le COS de manière conventionnelle, surtout en termes de stock, est donc fastidieux.

De plus, le COS est le carbone organique *total* du sol, qui est présent sous forme de constituants aux caractéristiques variées en termes de composition, dynamique et fonctions (Lehmann et Kleber, 2015). Plusieurs types de procédures ont été développés pour séparer le COS en fractions aux caractéristiques distinctes, et les approches de fractionnement physique se sont en particulier avérées fructueuses pour mieux comprendre la dynamique du COS (Christensen, 1992). Ces fractionnements permettent notamment de distinguer les matières organiques légères et grossières (> 0,05 mm), majoritairement d'origine végétale, jeunes et facilement décomposables, et les matières organiques fines associées à la phase minérale (< 0,05 mm), majoritairement d'origine microbienne, plus anciennes et moins facilement décomposables (Christensen, 1992 ; Lehmann et Kleber, 2015). Ces approches sont basées sur des technologies relativement simples mais fastidieuses à mettre en œuvre car elles nécessitent plusieurs étapes souvent longues et peu automatisables : dispersion chimique (sans altération des constituants organiques), extraction densimétrique éventuelle de la fraction légère, tamisages avec des tamis de maille décroissante, souvent ultrasonication pour disperser la fraction < 0,05 mm, tamisage encore, séchage et pesée de chaque fraction puis dosage du COS qu'elle contient. Malgré le grand intérêt de ces approches, leur lourdeur ne permet pas de les appliquer à un grand nombre d'échantillons. Il en est de même pour d'autres types de fractions organiques pouvant être distinguées dans le sol, par exemple la biomasse microbienne ou le COS potentiellement minéralisable.

Dans les décennies récentes, un intérêt croissant a été porté aux méthodes de caractérisation à haut débit et faible coût par échantillon, pour augmenter le nombre d'échantillons caractérisés. C'est dans ce contexte que l'utilisation de la spectroscopie dans l'infrarouge (IR) s'est développée pour caractériser les propriétés des sols, et tout particulièrement la concentration en COS (Dalal et Henry, 1986 ; Sudduth et Hummel, 1991 ; McCarty *et al.*, 2002 ; Barthès *et al.*, 2006 ; Bellon-Maurel et McBratney, 2011), ainsi que différents types de fractions organiques, par exemple granulométriques ou microbiennes (Ludwig *et al.*, 2002 ; Terhoeven-Urselmans *et al.*, 2006 ; Barthès *et al.*, 2008). En effet, c'est une approche rapide, à faible coût unitaire, sans consommable ni déchet, non destructive et reproductible ; de plus, diverses propriétés peuvent être caractérisées à partir d'un seul spectre. O'Rourke et Holden (2011) ont calculé que la détermination de la concentration en

COS d'un échantillon coûtait de l'ordre de dix fois moins cher par spectroscopie infrarouge que par combustion sèche. La plupart des usages actuels de la spectroscopie IR sont basés sur des calibrations, ou étalonnages, construits par régressions multivariées sur des échantillons de calibration, ou étalons, qui sont caractérisés à la fois conventionnellement (par exemple par combustion sèche pour le COS) et spectralement. Ces étalonnages peuvent ensuite être appliqués sur de nouveaux échantillons pour prédire la propriété considérée d'après leur spectre. En effet, la propriété n'est pas mesurée par spectroscopie IR mais prédite, avec une incertitude définie.

Cet article propose un aperçu d'ensemble montrant l'intérêt de la spectroscopie IR pour quantifier rapidement et à faible coût le carbone organique du sol et ses fractions.

2. Quantification de la concentration en carbone organique total du sol

2.1. Bases physiques et perspective historique

Le rayonnement IR provoque la vibration des liaisons moléculaires, et ces vibrations absorbent l'énergie lumineuse à des longueurs d'ondes qui dépendent de la force de la liaison, des atomes impliqués, et de l'environnement chimique (Siesler, 2007). Aussi, comparée à la lumière incidente, la lumière IR réfléchiée par un échantillon solide se trouve modifiée dans les régions spectrales correspondant à l'absorption d'énergie par les liaisons chimiques de l'échantillon. De manière générale, les vibrations fondamentales se produisent dans la région du moyen IR (MIR, $4000-400\text{ cm}^{-1}$, c'est-à-dire 2500 à 25 000 nm), tandis que leurs harmoniques et combinaisons vibrent dans la région proche IR (PIR, 800-2500 nm), où elles se chevauchent largement (Bellon-Maurel et McBratney, 2011). En conséquence, les spectres réfléchis présentent des pics beaucoup plus nets dans le MIR que dans le PIR (Stenberg *et al.*, 2010 ; cf. Figure 1) ; ce qui explique que le MIR (souvent appelé simplement « infrarouge ») ait été longtemps utilisé pour l'identification de molécules par reconnaissance de pics, sur la base de l'observation visuelle des spectres. Par exemple, les régions autour de $2750-2670$ et $1730-1710\text{ cm}^{-1}$ sont assignées aux acides carboxyliques aliphatiques saturés, et la région autour de 1650 cm^{-1} aux amides (Socrates, 2001 ; Barthès *et al.*, 2016).

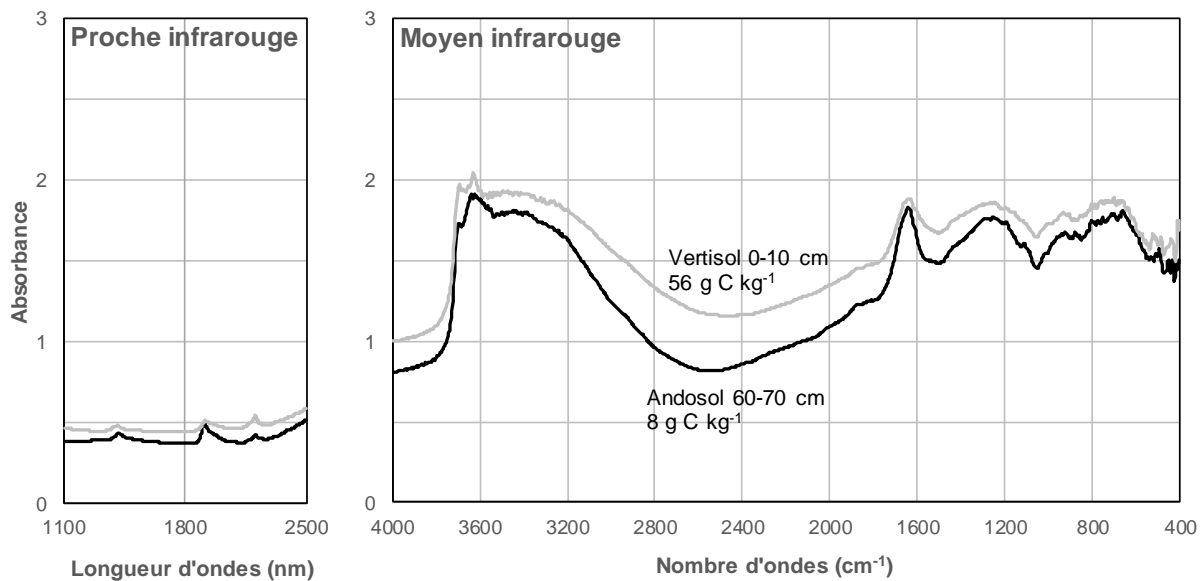


Figure 1 : Spectres proche et moyen infrarouge (IR) de deux échantillons de sols de la Martinique. L'abscisse est usuellement la longueur d'ondes (en nm) pour le proche IR et le nombre d'ondes (en cm^{-1}) pour le moyen IR ($2500 \text{ nm} = 4000 \text{ cm}^{-1}$). L'absorbance est plus forte et les pics plus informatifs dans le moyen que dans le proche IR (où les principaux pics, à 1400 et 1900 nm, sont assignés à l'eau).

Figure 1: Near and mid-infrared spectra of two soil samples from the island of Martinique. Abscissas are usually wavelengths (in nm) for near infrared and wavenumbers (in cm^{-1}) for mid-infrared ($2500 \text{ nm} = 4000 \text{ cm}^{-1}$). Absorbance is stronger and spectra display more informative peaks in the mid- than in the near infrared (where main peaks, at 1400 and 1900 nm, have been assigned to water).

Ces aspects de caractérisation qualitative voire semi-quantitative du COS par l'étude spécifique des pics ne seront pas abordés ici, parce que ce type d'approche ne permet pas le haut débit, ni de quantifier aisément le COS total. La reconnaissance de pics est beaucoup plus difficile sur les spectres PIR (sauf pour l'eau, cf. Figure 1), bien que différentes régions spectrales soient également assignées à des composés chimiques (Workman et Weyer, 2008 ; Barthès *et al.*, 2016). Des outils statistiques sont généralement requis pour exploiter les spectres PIR, d'où un développement plus récent des applications dans ce domaine spectral. Toutefois, l'absorption dans le MIR (vibrations fondamentales) étant beaucoup plus forte que dans le PIR (harmoniques et combinaisons), le rayonnement réfléchi est plus fort en PIR, d'où la moindre nécessité de capteurs très sensibles. Cela explique le développement

rapide des applications quantitatives de spectroscopie PIR, dès que furent largement disponibles les outils informatiques et d'analyse des données ; après un temps, les applications quantitatives en spectroscopie MIR ont suivi.

Cet article ne traite que de spectroscopie en réflexion diffuse, dans laquelle le capteur est placé dans l'une des multiples directions que prend la lumière réfléchiée par l'échantillon, car c'est de loin l'approche la plus commune en spectroscopie IR. Deux approches plus confidentielles utilisant le domaine MIR sont également explorées pour quantifier la concentration en COS : la réflexion totale atténuée (ATR en anglais), basée sur la radiation réfléchiée par les surfaces internes d'un cristal sur lequel l'échantillon est placé en contact étroit ; et la spectroscopie photo-acoustique, basée sur les ondes sonores créées lorsqu'un échantillon est chauffé par un rayonnement MIR (Reeves, 2010). Par ailleurs, outre le proche et le moyen infrarouge, le domaine infrarouge comprend aussi l'infrarouge lointain (25-350 μm), qui pour le moment connaît peu d'applications dans les domaines de l'environnement et de l'agronomie, notamment en science du sol. Notre propos sur l'infrarouge se limite donc au proche et au moyen infrarouge.

De très nombreux auteurs ont rapporté l'intérêt de la spectroscopie IR en réflexion diffuse pour caractériser les propriétés du sol, en particulier la concentration en COS, le plus souvent sur échantillons séchés à l'air et tamisés à 2 mm. Les premiers essais de prédiction de la concentration en COS à partir de spectres PIR de sol ont été réalisés par régression linéaire multiple, en utilisant la réflectance à quelques longueurs d'ondes (Dalal et Henry, 1986). Les étalonnages ont ensuite eu recours à des méthodes plus sophistiquées, régression en composantes principales puis régression des moindres carrés partiels (*partial least squares*, PLS ; Sudduth et Hummel, 1991) ; ces approches condensent l'information redondante des spectres (réflectance à plusieurs centaines ou milliers de longueurs d'ondes) en cinq à 25 variables dites latentes, qui sont des combinaisons linéaires de réflectances, orthogonales entre elles (indépendantes), et sont utilisées pour la régression. La régression PLS a ensuite été utilisée pour prédire la concentration en COS à partir de spectres MIR (Janik et Skjemstad, 1995). Depuis, de très nombreux travaux ont appliqué la régression PLS pour prédire la concentration en COS à partir de spectres VPIR (visible et proche infrarouge, 350-2500 nm ; c'est surtout la région PIR qui est utile), PIR ou MIR ; mais des procédures de régression non linéaire sont également utilisées, comme les régressions PLS dites locales (étalonnage à partir de voisins spectraux), et des approches de fouille de données ou

d'apprentissage automatique (*data mining, machine learning*) telles que les arbres de régressions renforcés (*boosted regression trees*), les forêts d'arbres de décision (*random forest*), les réseaux de neurones, *etc.* (Stenberg *et al.*, 2010 ; Ramirez-Lopez *et al.*, 2013 ; Vågen *et al.*, 2016). D'autres auteurs ont testé la procédure consistant à enrichir la base d'étalonnage avec quelques échantillons du site cible (*spiking* ; Guerrero *et al.*, 2014).

2.2. Evaluation de la qualité des prédictions

La qualité de l'ajustement entre observations et prédictions par spectroscopie IR est souvent évaluée par le coefficient de détermination R^2 , qui est un paramètre très classique, et dans une certaine mesure, permet des comparaisons entre études ; mais selon la manière dont il est calculé, il peut exprimer la proportionnalité plutôt que la similitude, et son utilisation peut donc conduire à des évaluations erronées. La racine de l'erreur quadratique moyenne (*root mean square error, RMSE*), souvent appelée erreur standard de prédiction, est plus précise mais ne permet guère de comparaisons si elle ne réfère pas à la distribution de la variable considérée (une erreur de prédiction de 3 gC.kg^{-1} est faible pour des sols riches en COS, mais élevée pour des sols pauvres en COS) ; de plus, ce paramètre permet difficilement des comparaisons entre études. Le RPD (ratio de la performance à la déviation) est le rapport entre écart type des valeurs observées et erreur standard de prédiction, qui intègre donc de l'information sur la distribution de la variable observée ; de ce fait, ce paramètre peut être utilisé pour comparer les prédictions de différentes variables, ou entre différentes études. Bellon-Maurel *et al.* (2010) ont toutefois considéré le RPD inapproprié pour les variables dont la distribution n'est pas normale (car l'écart type n'a alors guère de sens), ce qui est souvent le cas pour la concentration en COS ; ces auteurs ont proposé d'utiliser le RPIQ (ratio de la performance à la distance interquartile), qui est le rapport entre la distance interquartile des valeurs observées (entre troisième et premier quartile) et l'erreur standard de prédiction. Toutefois, RPIQ a encore peu été utilisé dans la littérature, aussi la suite du texte se focalisera-t-elle sur RPD et R^2 . Suite à Chang *et al.* (2001), de nombreux travaux considèrent que la prédiction de propriétés du sol par spectroscopie IR est précise quand $\text{RPD} \geq 2$, et acceptable quand $1.4 \leq \text{RPD} < 2.0$.

Comparer les performances des prédictions entre différentes études (ou variables) reste toutefois délicat, car plusieurs facteurs pas toujours simples à évaluer interviennent, en particulier :

- la diversité des jeux d'échantillons : les prédictions tendent à être moins précises quand la diversité texturale et minéralogique des échantillons augmente (cf. Brunet *et al.*, 2007), mais une certaine diversité de la variable à prédire est nécessaire pour étalonner (cf. van Groenigen *et al.*, 2003) ; *a contrario*, sur un jeu d'échantillons où la variable étudiée prend une gamme de valeurs relativement étroite, RPD (avec l'écart type au numérateur) et R^2 peuvent suggérer que la prédiction est médiocre, alors que l'erreur de prédiction est faible en valeur absolue, et certainement satisfaisante (cas par exemple du stock de COS sur une parcelle, cf. Figure 2) ;

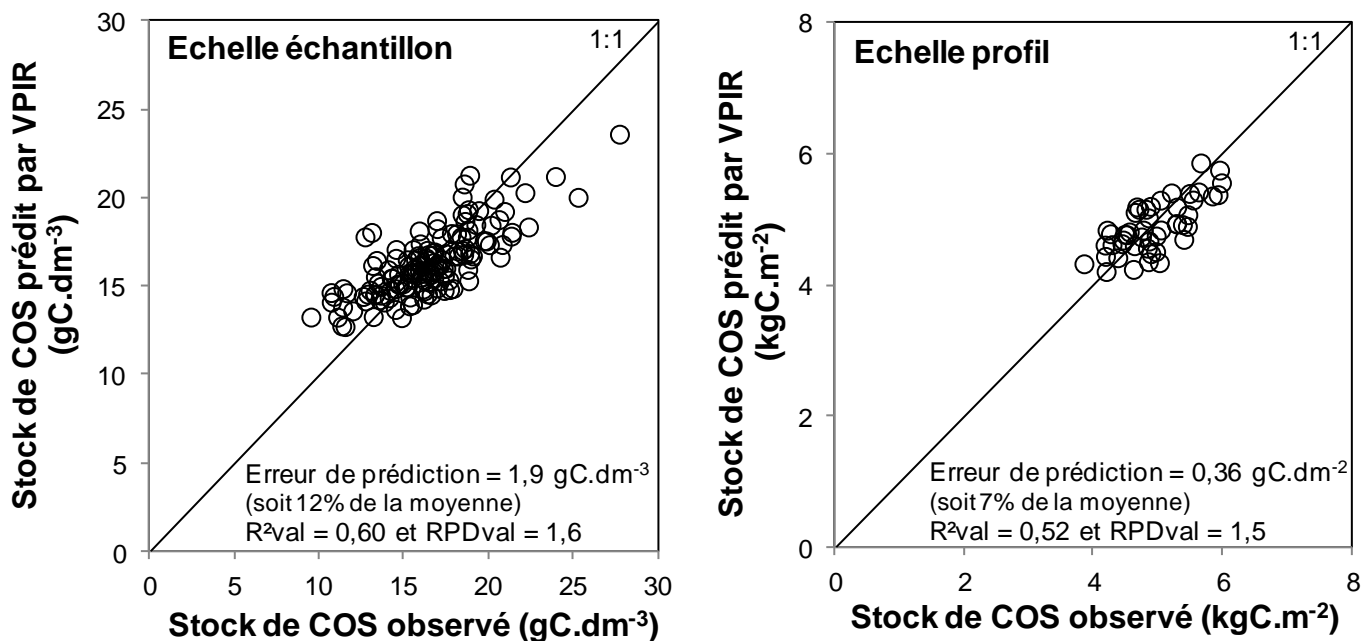


Figure 2 : Comparaison entre détermination conventionnelle (« observé ») et prédiction, à partir de spectres visible et proche infrarouge (VPIR) sur sol sec < 2 mm, du stock de carbone organique du sol (COS) à l'échelle de l'échantillon (gC.dm⁻³ ; à gauche) et du profil (kgC.m⁻² pour la couche 0-30 cm de profondeur ; à droite), sur un site de validation indépendant (distant de 250 à 700 km des sites d'étalonnage ; Cambou *et al.*, 2021).

Figure 2: Comparison between conventional determination (“observé”) and prediction, using visible and near infrared (“VPIR”) spectra on dry soil < 2 mm, of soil organic carbon (“COS”) stock at sample level (gC dm⁻³; on the left) and at profile level (kgC m⁻² for the 0-30 cm depth layer; on the right), for an independent validation site (250 to 700 km away from calibration sites; Cambou *et al.*, 2021).

- la robustesse des résultats de prédiction, c'est-à-dire l'étendue de leur domaine de validité : une validation externe indépendante (sur des échantillons ne provenant pas des sites d'étalonnage ; Brown *et al.*, 2005) donne souvent des résultats beaucoup plus robustes qu'une validation croisée « *leave-one-out* » (où chaque échantillon est prédit à tour de rôle à partir de tous les autres, ce qui donne des résultats optimistes quand certains échantillons proviennent des mêmes sites) ; de même, une validation croisée par groupes (où les échantillons de chaque groupe sont prédits à partir de ceux des autres groupes) donne des résultats plus robustes si tous les échantillons d'un site ou d'un profil sont dans un même groupe que s'ils se partagent entre groupes ; et quand un jeu d'échantillons est séparé en sous-jeu d'étalonnage et sous-jeu de validation externe, cette validation donne des résultats plus robustes si l'étalonnage est construit avec des échantillons tirés au hasard qu'avec les plus représentatifs (ce dernier cas est toutefois pertinent quand une grosse quantité d'échantillons est à caractériser, et qu'on en analyse une partie conventionnellement pour prédire l'autre partie par spectroscopie IR) ;
- la proportion d'échantillons utilisés pour l'étalonnage (sur un jeu fini d'échantillons) : plus la proportion d'étalons augmente, plus la précision en validation s'améliore, avec en général une atténuation aux proportions élevées (*cf.* Grinand *et al.*, 2012, qui ont utilisé 10 à 80 % d'une collection de plus de 2000 échantillons représentative des sols superficiels de France métropolitaine pour étalonner la concentration en COS puis la prédire sur les échantillons restant : la précision de la prédiction augmente fortement jusqu'à 20 % d'étalons, puis faiblement jusqu'à 40 %, et imperceptiblement au-delà).

Dans la suite du texte, R^2_{val} et RPD_{val} réfèrent à des prédictions obtenues en validation externe, tandis que R^2_{cv} et RPD_{cv} réfèrent à des prédictions obtenues en validation croisée.

2.3. Grosses bibliothèques spectrales (> 1000 échantillons)

Dans leur article de synthèse sur la quantification du COS par spectroscopie IR, Bellon-Maurel et McBratney (2011) ont insisté sur la nécessité de bases d'étalonnage, ou bibliothèques spectrales, aussi exhaustives que possibles, susceptibles de permettre des prédictions sur tous types d'échantillons de sols d'un territoire, quels que soient leurs minéralogie, texture, profondeur, usage, *etc.* De premières bibliothèques VPIR avaient été établies par Shepherd et Walsh (2002) pour l'Afrique orientale et australe, par Brown *et al.* (2006) principalement pour les Etats-Unis, et par Viscarra Rossel et Webster (2012) pour l'Australie (Tableau

supplémentaire 1). Ces librairies réunissaient des échantillons provenant de différents types d'expérimentations et de prospections. Des bases de données spectrales plus systématiques ont ensuite été constituées, dans la perspective de couvrir régulièrement des territoires pertinents d'un point de vue géographique ou administratif. Un travail notable a été réalisé en France avec les échantillons du Réseau de mesures de la qualité des sols (RMQS ; Martin *et al.*, 2022), prélevés à maille régulière (16 km × 16 km) sur tout le territoire métropolitain à deux niveaux de profondeur, et scannés en VPIR, PIR et MIR (Grinand *et al.*, 2012 ; Gogé *et al.*, 2012, 2014 ; Clairotte *et al.*, 2016). Plusieurs auteurs ont utilisé la librairie VPIR adossée au réseau LUCAS (*Land use/cover area frame statistical survey* ; Tóth *et al.*, 2013), qui représente les sols de surface de 23 pays européens, plus ou moins densément échantillonnés (1,1 à 7,7 échantillons par 1000 km² selon le pays ; Stevens *et al.*, 2013 ; Nocita *et al.*, 2014). Une grosse librairie VPIR a également été constituée en Chine (Shi *et al.*, 2015). Une librairie VPIR globale a même été construite à partir de spectres et d'analyses conventionnelles provenant de 92 pays, fournis par plusieurs dizaines de scientifiques (Viscarra Rossel *et al.*, 2016a) ; malgré la diversité des méthodes d'analyse du COS selon les échantillons (combustion sèche, oxydation sulfochromique, perte au feu, *etc.*), des prédictions satisfaisantes ont pu être obtenues avec un algorithme d'apprentissage automatique (*machine learning*), après pré-classification spectrale des échantillons.

La plupart des études utilisant de grosses librairies spectrales ont obtenu des prédictions précises de la concentration en COS, avec en général $R^2_{val} \approx 0,8$ et $RPD_{val} > 2$, et parfois $\approx 0,9$ et > 3 , respectivement, même avec la régression PLS usuelle (Tableau supplémentaire 1). Lorsque les procédures étaient optimisées, notamment la représentativité des étalons (voisins spectraux) et la gamme spectrale (par exemple, le MIR pour les sols des régions tempérées, *cf.* plus bas la section comparant PIR et MIR), l'erreur standard de prédiction se rapprochait de l'incertitude sur la méthode conventionnelle (par exemple 2,0 vs. 1,0–1,5 gC.kg⁻¹, respectivement, d'après Clairotte *et al.*, 2016, et Stevens *et al.*, 2013).

Compte tenu d'enjeux importants, plusieurs travaux récents proposent des stratégies pour optimiser l'utilisation des grosses librairies spectrales de sols, à l'échelle nationale par exemple, pour faire des prédictions à l'échelle locale (Barthès *et al.*, 2020 ; Ng *et al.*, 2021 ; Dorantes *et al.*, 2022 ; Li *et al.*, 2022 ; Shen *et al.*, 2022). Ces stratégies recourent souvent à l'enrichissement de la librairie avec quelques échantillons des sites cibles (« *spiking* »), et/ou

à des étalonnages donnant plus de poids aux étalons proches spectralement des échantillons cibles (étalonnages dits « locaux »).

2.4. Types d'échantillons de sol utilisés pour l'acquisition des spectres (granulométrie, humidité)

Au laboratoire, les spectres sont généralement acquis sur sol séché à l'air puis tamisé à 2 mm (terre fine), qui est le mode de préparation des échantillons de sol le plus commun pour leur analyse physico-chimique et leur conservation (Pansu *et al.*, 2001). Certaines analyses requièrent toutefois des préparations plus fines, notamment le dosage du carbone par combustion sèche (< 0,2 mm fréquemment ; Pansu et Gautheyrou, 2006). Des études comparant des prédictions à partir de spectres PIR ou VPIR de sol sec plus ou moins broyé ont montré qu'en général une préparation fine (de < 0,2 à < 0,5 mm) permet une meilleure prédiction de la concentration en COS qu'une préparation grossière (de < 2 à < 4 mm) pour les échantillons argileux, mais pas pour les échantillons à texture grossière ni pour des jeux d'échantillons de texture diversifiée (Dalal et Henry, 1986 ; Fystro, 2002 ; Reeves *et al.*, 2002 ; Russell, 2003 ; Barthès *et al.*, 2006 ; Brunet *et al.*, 2007 ; Barthès *et al.*, 2016 ; Hutengs *et al.*, 2019). Ces effets contradictoires selon la texture ont été expliqués par la taille et l'arrangement des particules, qui affectent la transmission de la lumière (Chang *et al.*, 2001 ; Barthès *et al.*, 2006).

Plusieurs auteurs ont même rapporté des prédictions VPIR plus précises de la concentration en COS à partir de spectres acquis sur carottes (décimétriques) ou mottes (centimétriques) séchées à l'air qu'à partir de spectres acquis sur sol séché à l'air < 2 mm, peut-être parce que la densité plus élevée des carottes et mottes permettrait une plus forte réflectance (Morgan *et al.*, 2009 ; Fontán *et al.*, 2010). Ce type de résultat ne semble toutefois possible que si la teneur en éléments grossiers (> 2 mm) est faible, ou peu variable ; sinon, il est difficile d'ajuster des spectres de sol total (à teneur élevée et variable en éléments grossiers) sur les concentrations en COS de la terre fine (< 2 mm).

Il s'est agi jusqu'ici d'acquisition de spectres sur sol sec. Stenberg (2010) a montré que la prédiction de la concentration en COS à partir de spectres VPIR de sol tamisé à 2 mm était plus précise lorsque les échantillons étaient réhumectés de manière homogène que lorsqu'ils étaient secs. En effet, l'humidité plus élevée réduit la différence de réfraction entre particules et pores, donc augmente la diffusion de la lumière vers l'avant (prodifusion), et

ainsi, la probabilité d'absorption d'un photon par une particule de sol (Twomey *et al.*, 1986). Toutefois, l'humectation homogène des échantillons peut difficilement être envisagée pour des acquisitions en routine de spectres au laboratoire. Sur le terrain, les échantillons sont souvent humides, mais l'humidité varie généralement entre échantillons, amenant de l'information perturbante dans les spectres, qui gêne les prédictions (car le spectre change avec l'humidité, mais la concentration en COS ne change pas). Deux études ont comparé des prédictions de la concentration en COS à partir de spectres VPIR acquis sur les mêmes carottes ou mottes, à l'humidité du terrain puis après séchage à l'air : elles montrent que par rapport aux prédictions sur sol séché à l'air, les prédictions à l'humidité du terrain sont plus précises si l'humidité varie peu entre échantillons (Gras *et al.*, 2014 ; spectres de terrain acquis à 0-20 cm en conditions hivernales assez humides), mais moins précises quand l'humidité varie notablement entre échantillons (Morgan *et al.*, 2009 ; spectres acquis à plusieurs profondeurs en saison chaude, parfois après orage ou irrigation). L'humidité améliore la transmission de la lumière, mais gêne les prédictions quand elle varie entre échantillons.

En spectroscopie MIR sur sol sec, la prédiction de la concentration en COS est nettement plus précise sur sol finement broyé (< 0,2 mm ou plus fin) que sur sol tamisé à 2 mm (Barthès *et al.*, 2016 ; Hutengs *et al.*, 2019 ; Wijewardane *et al.*, 2021). Le broyage fin permet de diminuer l'absorbance, très forte dans le MIR (vibrations fondamentales des liaisons), et donc d'augmenter la part du signal réfléchi vers le capteur ; de plus, la largeur du faisceau MIR étant de 1 à 2 mm, seul un broyage fin permet de scanner une partie représentative de l'échantillon (Reeves, 2010). La plupart des travaux de spectroscopie MIR sur sol sont d'ailleurs réalisés sur sol finement broyé. Cette recommandation du broyage fin pour le MIR est un inconvénient dans un contexte de caractérisation à haut débit.

2.5. Acquisition des spectres au laboratoire ou sur le terrain

La perspective d'acquérir les spectres directement sur le terrain suscite beaucoup d'intérêt, parce qu'alors sont supprimées les étapes de conditionnement, transport, séchage, concassage et tamisage à 2 mm (voire broyage fin) des échantillons, sauf pour les étalonnages. De très nombreux échantillons peuvent ainsi être caractérisés, augmentant encore l'efficacité de l'approche en matière de temps et de coût (Viscarra Rossel *et al.*, 2011). Toutefois, la prédiction de propriétés du sol à partir de spectres acquis sur le terrain

requiert, jusqu'à présent, des étalonnages également construits avec des spectres acquis sur le terrain, ce qui s'avère souvent fastidieux : alors qu'il existe des collections accessibles d'échantillons de sol séchés à l'air, tamisés à 2 mm et déjà analysés conventionnellement, à partir desquels la construction d'étalonnages est rapide, il n'existe pas de collections équivalentes d'échantillons déjà analysés disponibles dans l'état où ils sont scannés sur le terrain (structure, humidité). Donc construire des étalonnages pour des spectres de terrain nécessite des campagnes spécifiques de prélèvement et d'analyses conventionnelles, ce qui est particulièrement fastidieux et coûteux si la base d'étalonnage doit représenter un territoire vaste (par exemple un pays). Cette contrainte explique que la spectroscopie IR sur sol soit beaucoup moins appliquée sur le terrain qu'au laboratoire.

Sudduth et Hummel (1993) ont été les premiers à mentionner l'acquisition de spectres PIR de sol sur le terrain. Depuis, de nombreux travaux ont rapporté des prédictions de la concentration en COS à l'échelle locale, régionale ou multirégionale, à partir de spectres VPIR acquis dans des trous de tarière, sur carottes fraîches, sur échantillons frais émiettés, sur faces de fosses, depuis la surface, depuis une hauteur de 1 m, ou « en ligne » avec une sonde de spectromètre fixée sur une dent de chisel, *etc.* Les prédictions étaient souvent satisfaisantes, quoique moins précises en général que les éventuelles prédictions homologues en laboratoire sur sol sec < 2 mm : en validation croisée ou externe, $R^2 = 0,71$ à $0,93$ vs. $0,75$ à $0,95$, et $RPD = 1,9$ à $3,7$ vs. $2,0$ à $4,6$, respectivement (Mouazen *et al.*, 2006 ; Kusumo *et al.*, 2008, 2011 ; Stevens *et al.*, 2008 ; Nocita *et al.*, 2011 ; Gras *et al.*, 2014 ; Cambou *et al.*, 2016 ; Nawar et Mouazen, 2018 ; Allory *et al.*, 2019 ; Pei *et al.*, 2019 ; Allo *et al.*, 2020 ; Greenberg *et al.*, 2021) ; d'autres publications ont rapporté des prédictions de terrain moins précises, $R^2_{val} = 0,64-0,65$ et $RPD_{val} = 1,5-1,7$ vs. $0,73-0,82$ et $1,7-2,3$, respectivement (Morgan *et al.*, 2009 ; Veum *et al.*, 2018 ; Nawar et Mouazen, 2019), voire mauvaises, $R^2_{val} = 0,27$ et $RPD_{val} = 1,2$ vs. $0,66$ et $1,3$ sur sol sec < 2 mm, respectivement (Hutengs *et al.*, 2019 ; Tableau supplémentaire 2). Les prédictions moins précises sur le terrain qu'au laboratoire peuvent notamment être expliquées par l'humidité variable sur le terrain, et parce que les analyses conventionnelles sont réalisées sur les mêmes échantillons que ceux scannés au laboratoire (par exemple terre fine de la couche de sol 0-10 cm), qui ne correspondent pas toujours à ce qui est scanné au champ (par exemple spectre acquis en surface, ou sur sol riche en éléments grossiers > 2 mm). Néanmoins, quelques publications ont rapporté des prédictions VPIR de la concentration en COS un peu plus précises sur le

terrain (acquisition des spectres à 1 m de hauteur, depuis la surface, sur carotte entière ou émiettée) qu'au laboratoire sur sol sec < 2 mm : $R^2_{cv} = 0,76-0,86$ vs. $0,75-0,80$ et $RPD_{cv} = 2,1-2,8$ vs. $2,0-2,3$, respectivement (Stevens *et al.*, 2008 ; Gras *et al.*, 2014). Par rapport au sol sec < 2 mm, les échantillons scannés au champ (sol en place, carottes, mottes) sont en effet plus denses, et quand ils sont humides, ont une cohésion plus élevée, conditions qui favorisent la transmission de la lumière au sein de l'échantillon ; si l'humidité varie peu entre échantillons, que la charge grossière est faible et que les échantillons rapportés au laboratoire représentent correctement ceux scannés au champ, les prédictions peuvent alors être meilleures en conditions de terrain que sur sol sec < 2 mm.

La spectroscopie MIR portative est une technologie récente, et peu d'études de terrain ont pour l'instant été réalisées avec cette gamme spectrale : outre quelques travaux exploratoires de portée assez limitée (mono-sites, étalonnage peu robuste ; Izaurrealde *et al.*, 2013 ; Ji *et al.*, 2016), on note surtout les travaux d'une équipe allemande, qui a notamment comparé les prédictions à partir de spectres MIR acquis au laboratoire sur sol sec et au champ à la surface du sol (après décapage de la litière ; Hutengs *et al.*, 2019 ; Greenberg *et al.*, 2021, 2022 ; Tableau supplémentaire 2). La précision des prédictions MIR de terrain est généralement du même ordre qu'en VPIR (la comparaison VPIR vs. MIR est spécifiquement abordée dans la section suivante) ; et dans les comparaisons entre spectroscopie MIR au laboratoire et au champ, les résultats sont toujours moins bons au champ (Hutengs *et al.*, 2019 ; Greenberg *et al.*, 2021, 2022 ; les prédictions avec spectres de laboratoire de Greenberg *et al.*, 2021, ne sont pas rapportées dans le Tableau supplémentaire 2 car elles n'utilisent pas les mêmes spectromètres qu'au champ). Les moins bons résultats au champ sont là encore expliqués par l'humidité variable entre échantillons, mais on peut évoquer aussi une correspondance imparfaite entre spectres acquis sur la surface du sol et échantillons utilisés pour les analyses conventionnelles (et scannés au laboratoire), prélevés à 0-5 ou 0-2 cm de profondeur. On a vu aussi que le broyage fin des échantillons, en diminuant l'absorption très forte du rayonnement MIR, permet d'accroître le faible signal réfléchi mesuré par le capteur (*cf.* 2.4) ; le séchage diminue lui aussi l'absorption par l'échantillon, et il est donc logique que les prédictions MIR soient moins précises lorsque les spectres sont acquis sur sol ni sec ni broyé (faible signal réfléchi arrivant au capteur).

Il serait évidemment très avantageux de pouvoir appliquer à des spectres acquis sur le terrain (sol plus ou moins humide, structuré, pierreux, *etc.*) des étalonnages construits avec

des spectres de laboratoire (sur sol sec < 2 mm). Cela permettrait d'utiliser des bibliothèques spectrales constituées facilement à partir d'archives (sol sec < 2 mm) déjà analysées pour faire des prédictions à partir de spectres obtenus facilement sur le terrain. Une approche utilisée par plusieurs auteurs consiste à localiser dans les spectres acquis *in situ* la variabilité liée à l'humidité, à la structure, *etc.*, et à utiliser ensuite l'information spectrale non corrélée à cette variabilité (« orthogonale ») pour l'étalonnage et la prédiction du COS. Ainsi, Ge *et al.* (2014) et Veum *et al.* (2018) ont appliqué à des spectres acquis sur le terrain (sur carottes ou dans des sondages, respectivement) des étalonnages de la concentration en COS construits au laboratoire sur échantillons secs < 2 mm : les prédictions étaient meilleures avec que sans « orthogonalisation » des spectres de terrain par rapport à l'humidité et à la structure ($R^2_{val} = 0,46-0,53$ vs. $0,23-0,49$ et $RPD_{val} = 1,3-1,7$ vs. $0,5-1,3$, respectivement), mais restaient sensiblement moins précises que les prédictions utilisant les spectres de sol sec < 2 mm ($R^2_{val} = 0,69-0,82$ et $RPD_{val} = 2,1-2,3$). De telles approches d'orthogonalisation supposent que les informations spectrales relatives au COS, d'une part, et relatives à l'humidité et à la structure, d'autre part, ne soient pas corrélées entre elles, ce qui est discutable (par exemple, la teneur en eau n'est pas indépendante de la concentration en COS). La présence d'information spectrale sur le COS corrélée à l'information spectrale sur l'humidité et la structure, donc écartée par orthogonalisation, pourrait expliquer que celle-ci ne permette qu'une amélioration limitée.

2.6. Gamme spectrale : (visible et) proche infrarouge ou moyen infrarouge

Dans le sous-titre ci-dessus, « visible » est mis entre parenthèses parce que dans les spectres VPIR, c'est l'information du domaine PIR qui est la plus utile pour l'étude du COS ; par conséquent, ce qui est dit des spectres PIR concerne également les spectres VPIR.

De nombreux travaux ont rapporté de meilleures prédictions de la concentration en COS à partir de spectres MIR que de spectres PIR ou VPIR en conditions de laboratoire (Reeves *et al.*, 2001 ; McCarty *et al.*, 2002 ; Reeves *et al.*, 2002 ; McCarty et Reeves, 2006 ; Viscarra Rossel *et al.*, 2006 ; Igne *et al.*, 2010 ; Dong *et al.*, 2011 ; Xie *et al.*, 2011 ; Hutengs *et al.*, 2019 ; Greenberg *et al.*, 2021). Ces travaux concernent des sols de régions tempérées (sauf Dong *et al.*, 2011) et des jeux d'échantillons assez homogènes en général. Mais d'autres auteurs ont observé une tendance moins nette (Madari *et al.*, 2005) voire inverse, avec de meilleures prédictions (V)PIR que MIR (Madari *et al.*, 2006 ; Ludwig *et al.*, 2008 ; Rabenarivo

et al., 2013 ; Barthès *et al.*, 2016 ; Allo *et al.*, 2020), pour les régions tropicales et subtropicales (sauf Ludwig *et al.*, 2008) et des jeux d'échantillons plus diversifiés.

La supériorité supposée du MIR a été attribuée à l'information plus riche qu'il contient (*cf.* Figure 1) : pics mieux définis, assignés aux vibrations fondamentales, tandis que le PIR est une région d'harmoniques et de combinaisons, avec de nombreux chevauchements (Reeves, 2010 ; Bellon-Maurel et McBratney, 2011). Toutefois, dans les sols riches en oxydes de fer et d'aluminium (donc en particulier dans les régions tropicales et subtropicales), l'information concernant le COS pourrait être en partie masquée dans le domaine MIR du fait du chevauchement de pics liés aux composés minéraux et organiques (par exemple, les oxydes métalliques absorbent souvent à 1020-970 et 1100-825 cm^{-1} , tandis que les polysaccharides et autres glucides absorbent à 1080-1030 et 1170-950 cm^{-1} ; Rabenarivo *et al.*, 2013). Cela expliquerait les meilleures prédictions de la concentration en COS à partir de spectres (V)PIR que de spectres MIR pour les sols des régions (sub)tropicales. Ce résultat doit encore être confirmé, notamment sur des jeux d'échantillons larges et diversifiés.

Deux études ont comparé les prédictions de la concentration en COS à partir de spectres MIR ou (V)PIR acquis sur sol sec < 2 mm ou finement broyé (< 0,2 mm ou 0,01 mm) : la première, sur des sols tunisiens, rapporte de meilleures prédictions en PIR qu'en MIR, l'avantage du PIR étant faible sur sol < 0,2 mm mais marqué sur sol < 2 mm (Barthès *et al.*, 2016) ; en revanche, la seconde étude, sur des sols allemands, obtient des prédictions nettement meilleures en MIR qu'en VPIR, l'avantage du MIR étant plus marqué sur sol < 0,01 mm que sur sol < 2 mm (Hutengs *et al.*, 2019). Pour ce qui concerne la spectroscopie au laboratoire, ces résultats confirment l'intérêt du MIR pour les sols des régions tempérées et sur échantillons finement préparés, et l'intérêt du (V)PIR pour les sols des régions chaudes et sur échantillons < 2 mm. La recommandation d'acquérir les spectres MIR sur sol finement broyé est un inconvénient par rapport au PIR, dans un contexte de haut débit.

Deux comparaisons sur des sols allemands ont également été réalisées entre prédictions de la concentration en COS à partir de spectres MIR ou VPIR au laboratoire (sol sec finement broyé) ou au champ (spectres acquis sur la surface du sol) : au laboratoire, les prédictions étaient toujours plus précises en MIR qu'en VPIR, mais au champ, une étude rapporte de meilleures prédictions en MIR (Hutengs *et al.*, 2019) et l'autre, de meilleures prédictions en VPIR (Greenberg *et al.*, 2021 ; ces auteurs ont utilisé des spectromètres différents au

laboratoire et au champ, aussi leurs travaux de laboratoire ne sont-ils pas mentionnés dans le Tableau supplémentaire 2) ; ce qui ne permet pas de conclure.

3. Quantification du stock de carbone organique du sol

Le stock de COS d'un échantillon ($\text{gC} \cdot \text{dm}^{-3}$ sol total) est calculé comme le produit de sa concentration en COS ($\text{gC} \cdot \text{kg}^{-1}$ sol < 2 mm, le dosage étant effectué sur la terre fine) par sa masse volumique apparente (ou densité apparente, Da , $\text{kg sol total} \cdot \text{dm}^{-3}$ sol total), pondéré par la proportion massique de terre fine ($\text{kg sol} < 2 \text{ mm} \cdot \text{kg}^{-1}$ sol total ; Poeplau *et al.*, 2017), puis exprimée pour une couche de sol considérée ($\text{kgC} \cdot \text{m}^{-2}$ ou $\text{MgC} \cdot \text{ha}^{-1}$ pour cette couche de sol). Le stock est une variable additive, donc le stock à l'échelle du profil (par exemple à 0-30 cm de profondeur) est la somme des stocks de chaque échantillon considéré dans ce profil (par exemple 0-10, 10-20 et 20-30 cm), pour les déterminations conventionnelles comme pour les prédictions.

Certaines études ont estimé le stock de COS à partir de la mesure de Da et de la prédiction par spectroscopie IR de la concentration en COS : Cardinael *et al.* (2015) l'ont fait à partir de spectres VPIR acquis sur des carottes de sol intact à l'humidité du terrain, et Baldock *et al.* (2018) à partir de spectres MIR de sol séché à l'air et broyé finement ; mais ces travaux concernent la prédiction de la concentration en COS, non du stock de COS. Or, la détermination conventionnelle de Da est fastidieuse, car elle nécessite le prélèvement d'échantillons de volume apparent connu, ce qui requiert l'ouverture de fosses ou l'utilisation d'une carotteuse hydraulique mécanisée. Certains travaux ont utilisé d'autres approches pour caractériser Da . Sur une gamme de parcelles siciliennes, Priori *et al.* (2016), outre la prédiction de la concentration en COS à partir de spectres VPIR de sol sec < 2 mm, ont estimé Da avec une fonction de pédotransfert basée sur la texture et la concentration en COS mesurées ; ces auteurs ont prédit le stock de COS de la terre fine (< 2 mm ; $R^2_{\text{val}} = 0,77$ et $\text{RPD}_{\text{val}} = 2,1$), puis l'ont corrigé par une estimation semi-quantitative de la teneur en éléments grossiers (> 2 mm). Sur des carottes de sol intact à l'humidité du terrain, Lobsey et Viscarra Rossel (2016) et Viscarra Rossel *et al.* (2016b et 2017) ont utilisé la gammadensimétrie pour prédire Da , qui a ensuite été corrigée par l'humidité volumique prédite à partir des spectres VPIR, lesquels ont également servi à prédire la concentration en COS. Toutefois, ces travaux avaient pour objectif d'évaluer le stock de COS à l'échelle de la parcelle, mais ne présentent pas de comparaison entre stocks prédits et calculés

conventionnellement (sauf Viscarra Rossel *et al.*, 2016b, qui ont prédit le stock de COS à l'échelle du profil en utilisant le krigeage ; $R^2_{cv} = 0,4-0,5$ et $RPD_{cv} = 1,3-1,4$).

Les tentatives pour prédire D_a à partir de spectres IR ont rarement été satisfaisantes : plusieurs travaux ont obtenu de médiocres résultats ($R^2_{val} < 0,3$ et $RPD_{val} < 1,2$; Minasny *et al.*, 2008 ; Moreira *et al.*, 2009 ; Veum *et al.*, 2015). Certains auteurs ont rapporté de meilleurs résultats ($R^2_{val} = 0,64-0,81$ et $RPD_{val} = 1,7-2,2$; Askari *et al.*, 2015 ; Roudier *et al.*, 2015 ; Cambou *et al.*, 2016 ; Pérez-Fernández et Robertson, 2016) ; mais ce pourrait être dû à des jeux d'échantillons homogènes (Roudier *et al.*, 2015), ou au contraire à une très large distribution de D_a (l'erreur de prédiction étant en fait plus élevée que dans d'autres études ; Pérez-Fernández et Robertson, 2016), ou à l'agrégation de sites avec une médiocre prédiction intra-site (Cambou *et al.*, 2016).

Dans leur revue sur la quantification du stock de COS par spectroscopie IR, Bellon-Maurel et McBratney (2011) recommandaient de le prédire directement à partir des spectres (plutôt que de le calculer à partir de prédictions de la concentration et/ou de D_a), et considéraient comme prioritaire de développer des travaux dans ce sens. Suivant cette recommandation, plusieurs auteurs ont en effet obtenu des résultats prometteurs. Sur une grande parcelle cultivée, Roudier *et al.* (2015) ont fait des prédictions précises du stock de COS à l'échelle de l'échantillon (en $gC.dm^{-3}$) à partir de spectres VPIR acquis sur carottes intactes fraîches prélevées avec une carotteuse hydraulique mécanisée ($R^2_{val} = 0,85$ et $RPD_{val} = 2,6$). Sur deux parcelles cultivées distantes de 250 km, Cambou *et al.* (2016) ont obtenu une prédiction acceptable du stock de COS à l'échelle de l'échantillon à partir de spectres VPIR acquis sur carottes fraîches prélevées à la tarière à main ($R^2_{val} = 0,70$ et $RPD_{val} = 1,8$) ; ces carottes à structure perturbée contiennent moins d'information sur le stock que les carottes à structure intacte, d'où des prédictions moins précises que pour Roudier *et al.* (2015), imputables aussi à un jeu d'échantillons plus hétérogène. Sur une gamme de sites urbains contrastés, Allory *et al.* (2019) ont obtenu de bonnes prédictions du stock de COS à l'échelle de l'échantillon à partir de spectres VPIR acquis sur faces de fosse (structure intacte) ou sur sol sec < 2 mm ($R^2_{val} = 0,76-0,94$ et $RPD_{val} = 2,0-4,0$) ; et les prédictions tendaient à être plus précises pour le stock à l'échelle du profil (en $kgC.m^{-2}$; $R^2_{val} = 0,78-0,94$ et $RPD_{val} = 2,3-4,4$), par suite de compensations entre surestimations et sous-estimations à différents niveaux du profil. Reprenant trois études antérieures, Cambou *et al.* (2021) ont comparé le stock de COS à l'échelle de l'échantillon à partir de spectres VPIR de terrain (acquis

différemment selon l'étude) puis de sol sec < 2 mm : les prédictions se sont avérées plus précises à partir des spectres de sol sec < 2 mm qu'à partir des spectres de terrain, sauf sur un jeu d'échantillons à teneur variable en éléments grossiers > 2 mm (brouillant la relation entre sol < 2 mm et sol utilisé pour mesurer le stock). Les spectres VPIR de sol sec < 2 mm des trois études antérieures ont ensuite été réunis, avec un étalonnage utilisant tous les sites sauf un, et une validation indépendante sur le dernier site, donnant des prédictions satisfaisantes du stock de COS à l'échelle de l'échantillon et du profil 0-30 cm ($R^2_{val} = 0,52-0,60$ et $RPD_{val} = 1,5-1,6$; ces valeurs sont modestes car la gamme de stocks était étroite, mais l'erreur de prédiction, soit $0,36 \text{ kgC.m}^{-2}$, représentait 7 % du stock moyen, ce qui est satisfaisant ; Figure 2). A la Réunion, Allo *et al.* (2020) ont obtenu de très bonnes prédictions du stock de COS à l'échelle de l'échantillon à partir de spectres VPIR acquis sur carottes intactes fraîches ou sur sol sec < 2 mm ($R^2 = 0,87-0,91$ et $RPD = 2,7-3,3$ en validation croisée « *leave-one-out* ») ; les prédictions étaient moins précises à partir du spectre MIR de sol sec < 0,2 mm ($R^2 = 0,81$ et $RPD = 2,4$), confirmant la supériorité du VPIR sur le MIR pour quantifier le COS dans les régions tropicales (on peut également supposer que le broyage à 0,2 mm a détruit une partie de l'information sur le stock). En revanche, aucune publication n'a encore rapporté la prédiction du stock de COS par spectroscopie MIR de terrain. Ces différents travaux soulignent l'intérêt des spectres VPIR de sol sec < 2 mm (quand la pierrosité est faible) ou de terrain pour prédire directement le stock de COS.

Enfin, Cambou *et al.* (2021) ont montré que ce ne sont pas exactement les mêmes régions spectrales qui contribuent le plus à la prédiction de la concentration et du stock de COS, et que la prédiction du stock est parfois plus précise que celle de la concentration. Ces résultats réfutent l'hypothèse que la prédiction du stock découlerait de la prédiction de la concentration et de la corrélation entre concentration et stock.

4. Caractérisation de la qualité de la matière organique du sol

Cette section porte non plus sur la quantification du carbone organique *total* du sol, mais sur la caractérisation de la distribution granulométrique ou granulodensimétrique du COS (les fractionnements granulométriques sont basés seulement sur la taille des particules, tandis que les fractionnements granulodensimétriques comportent aussi une extraction de la fraction légère). Cette caractérisation est particulièrement utile pour mieux comprendre la dynamique et les fonctions du COS (Christensen, 1992).

Plusieurs travaux ont rapporté des prédictions de la distribution du COS dans les fractions granulo(densi)métriques à partir du spectre de la terre fine, en (V)PIR (Cozzolino et Morón, 2006 ; Barthès *et al.*, 2008 ; Yang *et al.*, 2012 ; Jaconi *et al.*, 2019 ; Greenberg *et al.*, 2021 ; Ramifehiarivo *et al.*, 2023) ou en MIR (Zimmermann *et al.*, 2007 ; Yang *et al.*, 2012 ; Baldock *et al.*, 2013, 2018 ; Greenberg *et al.*, 2021 ; Ramifehiarivo *et al.*, 2023). Seules seront considérées ici les prédictions réalisées au laboratoire, sur sol séché et tamisé à 2 mm voire broyé finement, car une seule étude a pour l'instant rapporté des prédictions à partir de spectres de terrain (depuis la surface ; Greenberg *et al.*, 2021). La qualité des prédictions est bonne en général, avec $R^2_{val} = 0,7-0,9$ et $RPD_{val} = 2-3$ le plus souvent (*cf.* par exemple Figure 3).

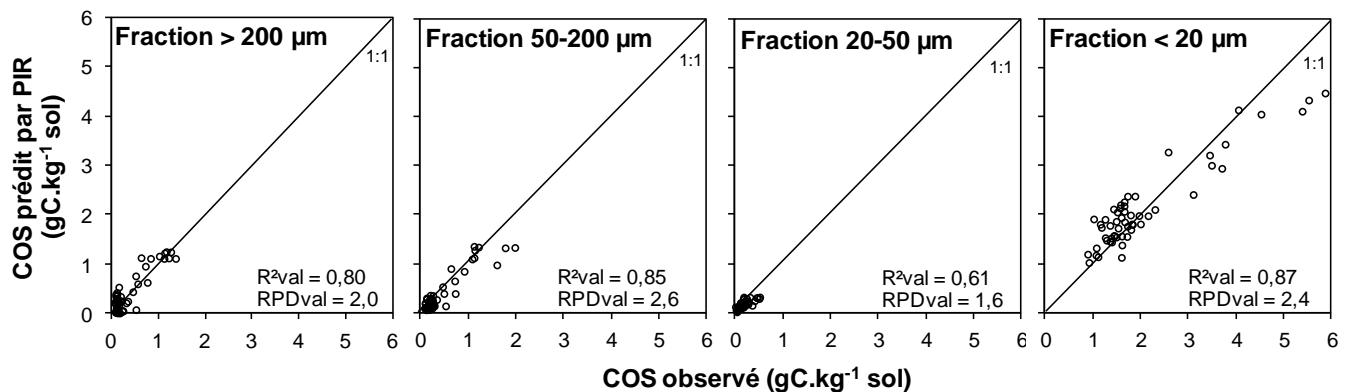


Figure 3 : Comparaison entre détermination conventionnelle (« observé ») et prédiction, à partir de spectres proche infrarouge (PIR) sur sol sec < 2 mm, de la distribution granulométrique du COS dans des sols du Burkina Faso (données de Barthès *et al.*, 2008, en excluant les sites du Congo-Brazzaville ; les 100 échantillons les plus représentatifs spectralement ont été utilisés pour l'étalonnage et les 55 restants pour la validation présentée ici).

Figure 3: Comparison between conventional determination (“observé”) and prediction, using near infrared (“PIR”) spectra of dry soil < 2 mm, of particle-size distribution of soil organic carbon (“COS”) in soils from Burkina Faso (data from Barthès *et al.*, 2008, after the removal of samples from Congo-Brazzaville; the 100 samples most representative spectrally were used for calibration, and the 55 remaining samples for the validation presented here).

Les prédictions sont parfois moins précises pour les fractions de faible masse (*cf.* Figure 3, $R^2_{val} = 0,61$ et $RPD_{val} = 1,6$ pour la fraction 20-50 µm, qui ne représente en moyenne que

10 % de la masse de la terre fine des sols considérés), ou plus précises pour les fractions représentant une proportion importante du sol (par exemple Ramifehiarivo *et al.*, 2023, obtiennent en validation externe non indépendante $R^2_{val} = 0,95$ et $RPD_{val} = 4,3-4,5$ pour $COS < 20 \mu m$ dans des sols où cette fraction représente en moyenne 46 % de la terre fine). Une étude comparant MIR et VPIR pour prédire de telles fractions sur des sols allemands ne permet pas de départager les deux gammes spectrales (Greenberg *et al.*, 2021), tandis qu'une étude sur sols malgaches montre un léger avantage du PIR sur le MIR (Ramifehiarivo *et al.*, 2023).

Les spectres MIR (et parfois VPIR) de laboratoire sont acquis sur sol finement broyé ($< 0,2 \text{ mm}$ ou plus fin), ce qui *a priori* devrait compliquer la prédiction du COS dans les fractions grossières ($> 200 \mu m$ en particulier) ; mais ce n'est pas ce qui est observé en général ($R^2_{val} = 0,77-0,97$ et $RPD_{val} = 2,0-3,4$ pour le COS particulaire ; Zimmermann *et al.*, 2007 ; Baldock *et al.*, 2013, 2018 ; Jaconi *et al.*, 2019 ; Ramifehiarivo *et al.*, 2023), ce qui témoigne que les fractions organiques grossières ont une composition spécifique (Christensen, 1992 ; Lehmann et Kleber, 2015). Calderón *et al.* (2011) ont effectivement montré que les différentes fractions granulodensimétriques du sol ont des spectres MIR contrastés.

D'autres types de fractions du COS peuvent également être caractérisés à partir du spectre (V)PIR ou MIR de la terre fine : COS potentiellement minéralisable (Ludwig *et al.*, 2002 ; Barthès *et al.*, 2016), carbone microbien (Ludwig *et al.*, 2002 ; Terhoeven-Urselmans *et al.*, 2006), fractions RMN (Terhoeven-Urselmans *et al.*, 2006), COS oxydable au permanganate (POXC ; Calderón *et al.* 2017), etc.

5. Conclusion

La spectroscopie IR offre de réelles opportunités pour caractériser le COS à haut débit et faible coût, et ainsi répondre aux enjeux de gestion durable des écosystèmes. Plusieurs gammes spectrales peuvent être utilisées : en général le MIR donne de meilleures prédictions que le PIR pour les sols des régions tempérées, à condition que les échantillons soient secs et finement broyés ($< 0,2 \text{ mm}$ ou plus fin) ; et l'inverse est rapporté pour les sols des régions (sub)tropicales ou lorsque les spectres sont acquis sur sol grossièrement préparé (par exemple $< 2 \text{ mm}$) ou *in situ* (avec les instruments actuels du moins).

L'acquisition des spectres sur le terrain (sur carottes, faces de fosse, *etc.*) est une perspective séduisante, car elle élimine les étapes de conditionnement, transport et préparation des échantillons (sauf pour les étalonnages). Toutefois, la construction d'étalonnages avec ces spectres nécessite jusqu'à présent des campagnes spécifiques de prélèvement et d'analyse de sol, et est donc sensiblement plus laborieuse et coûteuse qu'avec des spectres acquis au laboratoire sur sol sec tamisé ou broyé (car des collections d'échantillons préparés ainsi et déjà analysés sont assez facilement disponibles pour construire des étalonnages). Par ailleurs, les prédictions sont souvent moins précises à partir de spectres acquis sur le terrain que sur sol sec, du fait principalement des différences d'humidité entre échantillons. Toutefois, des approches sont explorées pour réduire l'effet de l'humidité et de la structure sur les spectres de terrain, dans la perspective d'appliquer à des spectres de terrain des étalonnages construits sur sol sec < 2 mm.

La spectroscopie IR a été largement utilisée depuis plusieurs décennies pour prédire la concentration en COS (gC.kg^{-1}), mais des travaux récents montrent qu'elle permet également de prédire le stock de COS (gC.dm^{-3} , ou kgC.m^{-2} pour une couche de sol donnée), sur le terrain ou au laboratoire. Les prédictions sont plus précises à partir de spectres VPIR acquis sur sol sec < 2 mm qu'à partir de spectres acquis sur le terrain, sauf quand la teneur en éléments grossiers (> 2 mm) est élevée et variable. Ces travaux ouvrent d'importantes perspectives car la caractérisation conventionnelle du stock de COS est fastidieuse, spécialement la détermination de la densité apparente, dont la spectroscopie VPIR permet de s'affranchir (sauf pour les étalonnages).

Par ailleurs, la spectroscopie IR a montré son intérêt pour quantifier à partir du spectre de la terre fine différents types de fractions du COS, en particulier pour caractériser sa distribution granulométrique ou granulodensimétrique, ce qui ouvre également d'importantes perspectives. En effet, caractériser les fractions du COS donne accès à des informations cruciales notamment pour mieux comprendre la dynamique de la matière organique du sol, mais est coûteux avec les méthodes conventionnelles.

Les applications de spectroscopie IR aux sols sont encore en phase de développement, en particulier pour rendre les bibliothèques spectrales aussi exhaustives qu'il est nécessaire (diversité des sols représentés) et interopérables (utilisation conjointe de spectres acquis avec différents instruments), mais aussi pour optimiser les procédures d'étalonnage. Toutefois, même lorsque cette phase de développement s'achèvera, la spectroscopie IR ne

se substituera pas aux méthodes conventionnelles d'analyse de sol, dont elle continuera à dépendre pour construire ou au moins compléter les étalonnages nécessaires aux prédictions. Mais elle permet de beaucoup mieux utiliser et valoriser les analyses conventionnelles : en effet, une analyse conventionnelle réalisée sur un échantillon ne sert plus seulement à caractériser cet échantillon particulier, mais, couplée à un spectre (rapide et peu coûteux à acquérir), elle contribue à des étalonnages permettant de caractériser de nombreux autres échantillons. L'ère des données rares et coûteuses sur les sols laisse place à une ère de données abondantes et peu coûteuses, un peu moins précises mais à l'imprécision définie.

6. Remerciements

Merci à Alain Albrecht d'avoir informé Christian Feller des travaux de spectroscopie proche IR sur les sols dont il avait eu connaissance au Kenya, à l'ICRAF (International Centre for Research in Agroforestry) ; merci à Christian Feller d'avoir compris l'intérêt de l'approche et d'avoir décidé d'équiper son unité de recherche Séquestration du carbone dans les sols tropicaux (UR SéqC, IRD) d'un premier spectromètre en 2002 ; merci à Didier Brunet d'avoir contribué au développement des applications de spectroscopie IR en science du sol à l'UR SéqC puis à l'UMR Eco&Sols (Ecologie fonctionnelle et biogéochimie des sols et agroécosystèmes ; Institut Agro Montpellier – Cirad – Inrae – IRD). Les deux relecteurs d'une première version du manuscrit sont également remerciés pour leur contribution à son amélioration.

7. Références

- Allo M., Todoroff P., Jameux M., Stern M., Paulin L., Albrecht A. (2020). Prediction of tropical volcanic soil organic carbon stocks by visible-near- and mid-infrared spectroscopy. *Catena*, 189, 104452. DOI: 10.1016/j.catena.2020.104452
- Allory V., Cambou A., Moulin P., Schwartz C., Cannavo P., Vidal-Beaudet L., Barthès B.G. (2019). Quantification of soil organic carbon stock in urban soils using visible and near infrared reflectance spectroscopy (VNIRS) *in situ* or in laboratory conditions. *Science of the Total Environment*, 686, pp. 764-773. DOI: 10.1016/j.scitotenv.2019.05.192.

- Askari M.S., O'Rourke S.M., Holden N.M. (2015). Evaluation of soil quality for agricultural production using visible-near-infrared spectroscopy. *Geoderma*, 243–244, pp. 80-91. DOI: 10.1016/j.geoderma.2014.12.012
- Baldock J.A., Hawke B., Sanderman J., Macdonald L.M. (2013). Predicting contents of carbon and its component fractions in Australian soils from diffuse reflectance mid-infrared spectra. *Soil Research*, 51, pp. 577-595. DOI: 10.1071/SR13077
- Baldock J.A., Beare M.H., Curtin D., Hawke B. (2018). Stocks, composition and vulnerability to loss of soil organic carbon predicted using mid-infrared spectroscopy. *Soil Research*, 56, pp. 464-480. DOI: 10.1071/SR17221
- Barthès B.G., Brunet D., Ferrer H., Chotte J.L., Feller C. (2006). Determination of total carbon and nitrogen content in a range of tropical soils using near infrared spectroscopy: Influence of replication and sample grinding and drying. *Journal of Near Infrared Spectroscopy*, 14, pp. 341-348. DOI: 10.1255/jnirs.686
- Barthès B.G., Brunet D., Hien E., Enjalric F., Conche S., Freschet G.T., d'Annunzio R., Toucet-Louri J. (2008). Determining the distributions of soil carbon and nitrogen in particle size fractions using near-infrared reflectance spectrum of bulk soil samples. *Soil Biology and Biochemistry*, 40, pp. 1533-1537. DOI: 10.1016/j.soilbio.2007.12.023
- Barthès B.G., Kouakoua E., Moulin P., Hmaid K., Gallali T., Clairotte M., Bernoux M., Bourdon E., Toucet J., Chevallier T. (2016). Studying the physical protection of soil carbon with quantitative infrared spectroscopy. *Journal of Near Infrared Spectroscopy*, 24, pp. 199-214. DOI: 10.1255/jnirs.1232
- Barthès B.G., Kouakoua E., Coll P., Clairotte M., Moulin P., Saby N.P.A., Le Cadre E., Etayo A., Chevallier T. (2020). Improvement in spectral library-based quantification of soil properties using representative spiking and local calibration – The case of soil inorganic carbon prediction by mid-infrared spectroscopy. *Geoderma*, 369, 114272. DOI: 10.1016/j.geoderma.2020.114272
- Bellon-Maurel V., Fernandez-Ahumada E., Palagos B., Roger J.M., McBratney A. (2010). Prediction of soil attributes by NIR spectroscopy. A critical review of chemometric indicators commonly used for assessing the quality of the prediction. *Trends in Analytical Chemistry*, 29, pp. 1073-1081. DOI: 10.1016/j.trac.2010.05.006.
- Bellon-Maurel V., McBratney A. (2011). Near-infrared (NIR) and mid-infrared (MIR) spectroscopic techniques for assessing the amount of carbon stock in soils – Critical review and research perspectives. *Soil Biology and Biochemistry*, 43, pp. 1398-1410. DOI: 10.1016/j.soilbio.2011.02.019.
- Brown D.J., Bricklemyer R.S., Miller P.R. (2005). Validation requirements for diffuse reflectance soil characterization models with a case study of VNIR soil C prediction in Montana. *Geoderma*, 129, pp. 251-267. DOI: 10.1016/j.geoderma.2005.01.001

- Brown D.J., Shepherd K.D., Walsh M.G., Mays M.D., Reinsch T.G. (2006). Global soil characterization with VNIR diffuse reflectance spectroscopy. *Geoderma*, 132, pp. 273–290. DOI: 10.1016/j.geoderma.2005.04.025
- Brunet D., Barthès B.G., Chotte J.L., Feller C. (2007). Determination of carbon and nitrogen contents in Alfisols, Oxisols and Ultisols from Africa and Brazil using NIRS analysis: Effects of sample grinding and set heterogeneity. *Geoderma*, 139, pp. 106-117. DOI: 10.1016/j.geoderma.2007.01.007
- Calderón F.J., Reeves J.B. III, Collins H.P., Paul E.A. (2011). Chemical differences in soil organic matter fractions determined by diffuse-reflectance mid-infrared spectroscopy. *Soil Science Society of America Journal*, 75, pp. 568-579. DOI: 10.2136/sssaj2009.0375
- Calderón F.J., Culman S., Six J., Franzluebbers A.J., Schipanski M., Beniston J., Grandy S., Kong A.Y.Y. (2017). Quantification of soil permanganate oxidizable C (POXC) using infrared spectroscopy. *Soil Science Society of America Journal*, 81, pp. 277-288. DOI: 10.2136/sssaj2016.07.0216
- Cambou A., Cardinael R., Kouakoua E., Villeneuve M., Durand C., Barthès B.G. (2016). Prediction of soil organic carbon stock using visible and near infrared reflectance spectroscopy (VNIRS) in the field. *Geoderma*, 261, pp. 151-159. DOI: 10.1016/j.geoderma.2015.07.007
- Cambou A., Allory V., Cardinael R., Carvalho Vieira L., Barthès B.G. (2021). Comparison of soil organic carbon stocks predicted using visible and near infrared reflectance (VNIR) spectra acquired *in situ* vs. on sieved dried samples: Synthesis of different studies. *Soil Security*, 5, 100024. DOI: 10.1016/j.soisec.2021.100024
- Cardinael R., Chevallier T., Barthès B.G., Saby N.P., Parent T., Dupraz C., Bernoux M., Chenu C. (2015). Impact of alley cropping agroforestry on stocks, forms and spatial distribution of soil organic carbon – A case study in a Mediterranean context. *Geoderma*, 259-260, pp. 288-299. DOI: 10.1016/j.geoderma.2015.06.015.
- Chang C.W., Laird D.A., Mausbach M.J., Hurburgh C.R. Jr (2001). Near-infrared reflectance spectroscopy–principal components regression analyses of soil properties. *Soil Science Society of America Journal*, 65, pp. 480-490. DOI: 10.2136/sssaj2001.652480x
- Christensen B.T. (1992). Physical fractionation of soil and organic matter in primary particle size and density separates. *Advances in Soil Science*, 20, pp. 1-90. DOI: 10.1007/978-3-540-31211-6_9
- Clairotte M., Grinand C., Kouakoua E., Thébault A., Saby N.P.A., Bernoux M., Barthès B.G. (2016). National calibration of soil organic carbon concentration using diffuse infrared reflectance spectroscopy. *Geoderma*, 276, pp. 41-52. DOI: 10.1016/j.geoderma.2016.04.021
- Cozzolino D., Morón A. (2006). Potential of near-infrared reflectance spectroscopy and chemometrics to predict soil organic carbon fractions. *Soil and Tillage Research*, 85, pp. 78-85. DOI: 10.1016/j.still.2004.12.006

- Dalal R.C., Henry R.J. (1986). Simultaneous determination of moisture, organic carbon, and total nitrogen by near infrared reflectance spectrophotometry. *Soil Science Society of America Journal*, 50, pp. 120-123. DOI: 10.2136/sssaj1986.03615995005000010023x
- Derrien D., Dignac M.F., Basile-Doelsch I., Barot S., Cécillon L., Chenu C., Chevallier T., Freschet G.T., Garnier P., Guenet B., Hedde M., Klumpp K., Lashermes G., Maron P.A., Nunan N., Roumet C., Barré P. (2016). Stocker du C dans les sols : quels mécanismes, quelles pratiques agricoles, quels indicateurs ? *Etude et Gestion des Sols*, 23, pp. 193-223. https://www.afes.fr/wp-content/uploads/2017/09/EGS_23_1_2314_derriendignac193_224_V3.pdf (consulté le 24/10/2022)
- Dong Y.W., Yang S.Q., Xu C.Y., Li Y.Z., Bai W., Fan Z.N., Wang Y.N., Li Q.Z. (2011). Determination of soil parameters in apple-growing regions by near- and mid-infrared spectroscopy. *Pedosphere*, 21, pp. 591-602. DOI: 10.1016/S1002-0160(11)60161-6
- Dorantes M.J., Fuentes B.A., Miller D.M. (2022). Calibration set optimization and library transfer for soil carbon estimation using soil spectroscopy – A review. *Soil Science Society of America Journal*, 86, pp. 879-903. DOI: 10.1002/saj2.20435
- Fontán J.M., Calvache S., López-Bellido R.J., López-Bellido L. (2010). Soil carbon measurement in clods and sieved samples in a Mediterranean Vertisol by Visible and Near-Infrared Reflectance Spectroscopy. *Geoderma*, 156, pp. 93-98. DOI: 10.1016/j.geoderma.2010.02.001
- Fystro G. (2002). The prediction of C and N content and their potential mineralisation in heterogeneous soil samples using Vis–NIR spectroscopy and comparative methods. *Plant and Soil*, 246, pp. 139-149. DOI: 10.1023/A:1020612319014
- Ge Y., Morgan C.L.S., Ackerson J.P. (2014). VisNIR spectra of dried ground soils predict properties of soils scanned moist and intact. *Geoderma*, 221-222, pp. 61-69. DOI: 10.1016/j.geoderma.2014.01.011
- GIEC (Groupe d'experts intergouvernemental sur l'évolution du climat) (2019). Résumé à l'intention des décideurs, *Changement climatique et terres émergées : rapport spécial du GIEC sur le changement climatique, la désertification, la dégradation des sols, la gestion durable des terres, la sécurité alimentaire et les flux de gaz à effet de serre dans les écosystèmes terrestres*. IPCC, Genève.
https://www.ipcc.ch/site/assets/uploads/sites/4/2020/06/SRCCL_SPM_fr.pdf (consulté le 28/09/2022).
- Gogé F., Joffre R., Jolivet C., Ross I., Ranjard L. (2012). Optimization criteria in sample selection step of local regression for quantitative analysis of large soil NIRS database. *Chemometrics and Intelligent Laboratory Systems*, 110, pp. 168-176. DOI: 10.1016/j.chemolab.2011.11.003

- Gogé F., Gomez C., Jolivet C., Joffre R. (2014). Which strategy is best to predict soil properties of a local site from a national Vis–NIR database? *Geoderma*, 213, pp. 1-9. DOI: j.geoderma.2013.07.016
- Gras J.P., Barthès B.G., Mahaut B., Trupin S. (2014). Best practices for obtaining and processing field visible and near infrared (VNIR) spectra of topsoils. *Geoderma*, 214-215, pp. 126-134. DOI: 10.1016/j.geoderma.2013.09.021
- Greenberg I., Seidel M., Vohland M., Ludwig B. (2021). Performance of field-scale lab vs in situ visible/near- and mid-infrared spectroscopy for estimation of soil properties. *European Journal of Soil Science*, 73, e13180. DOI: 10.1111/ejss.13180
- Greenberg I., Seidel M., Vohland M., Koch H.J., Ludwig B. (2022). Performance of in situ vs laboratory mid-infrared soil spectroscopy using local and regional calibration strategies. *Geoderma*, 409, 115614. DOI: 10.1016/j.geoderma.2021.115614
- Grinand C., Barthès B.G., Brunet D., Kouakoua E., Arrouays D., Jolivet C., Caria G., Bernoux M. (2012). Prediction of soil organic and inorganic carbon contents at a national scale (France) using mid-infrared reflectance spectroscopy (MIRS). *European Journal of Soil Science*, 63, pp. 141-151. DOI: 10.1111/j.1365-2389.2012.01429.x
- Guerrero C., Stenberg B., Wetterlind J., Viscarra Rossel R.A., Maestre F.T., Mouazen A.M., Zornoza R., Ruiz-Sinoga J.D., Kuang B. (2014). Assessment of soil organic carbon at local scale with spiked NIR calibrations: effects of selection and extra-weighting on the spiking subset. *European Journal of Soil Science*, 65, pp. 248-263. DOI: 10.1111/ejss.12129
- Hutengs C., Seidel M., Oertel F., Ludwig B., Vohland M. (2019). In situ and laboratory soil spectroscopy with portable visible-to-near infrared and mid-infrared instruments for the assessment of organic carbon in soils. *Geoderma*, 355, 113900. DOI: geoderma.2019.113900
- Igne B., Reeves J.B. III, McCarty G., Hively W.D., Lund E., Hurburgh C.R. Jr (2010). Evaluation of spectral pretreatments, partial least squares, least squares support vector machines and locally weighted regression for quantitative spectroscopic analysis of soils. *Journal of Near Infrared Spectroscopy*, 18, pp. 167-176. DOI: 10.1255/jnirs.883
- Initiative Internationale « 4 pour 1000 » (2018). <http://4p1000.org> (consulté le 24/10/2022).
- Izaurrealde R.C., Rice C.W., Wielopolski L., Ebinger M.H., Reeves J.B. III, Thomson A.M., Harris R., Francis B., Mitra S., Rappaport A.G., Etchevers J.D., Sayre K.D., Govaerts B., McCarty G.W. (2013). Evaluation of three field-based methods for quantifying soil carbon. *PLoS One*, 8, e55560. DOI: 10.1371/journal.pone.0055560
- Jaconi A., Poeplau C., Ramirez-Lopez L., Van Wesemael B., Don A. (2019). Log-ratio transformation is the key to determining soil organic carbon fractions with near-infrared spectroscopy. *European Journal of Soil Science*, 70, pp. 127-139. DOI: 10.1111/ejss.12761

- Janik L.J., Skjemstad J.O. (1995). Characterization and analysis of soils using mid-infrared partial least-squares. II. Correlations with some laboratory data. *Australian Journal of Soil Research*, 33, pp. 637-650. DOI: 10.1071/SR9950637
- Ji W., Adamchuk V.I., Biswas A., Dhawale N.M., Sudarsan B., Zhang Y., Viscarra Rossel R.A., Shi Z. (2016). Assessment of soil properties *in situ* using a prototype portable MIR spectrometer in two agricultural fields. *Biosystems Engineering*, 152, pp. 14-27. DOI: 10.1016/j.biosystemseng.2016.06.005
- Kusumo B.H., Hedley C.B., Hedley M.J., Hueni A., Tuohy M.P., Arnold G. (2008). The use of diffuse reflectance spectroscopy for *in situ* carbon and nitrogen analysis of pastoral soils. *Soil Research*, 46, pp. 623-635. DOI: 10.1071/SR08118
- Kusumo B.H., Hedley M.J., Hedley C.B., Tuohy M.P. (2011). Measuring carbon dynamics in field soils using soil spectral reflectance: prediction of maize root density, soil organic carbon and nitrogen content. *Plant and Soil*, 338, pp. 233-245. DOI: 10.1007/s11104-010-0501-4
- Lal R. (2014). Soil conservation and ecosystem services. *International Soil and Water Conservation Research*, 2, pp. 36-47. DOI: 10.1016/S2095-6339(15)30021-6
- Li H.Y., Li Y.H., Yang M.H., Chen S.C., Shi Z. (2022). Strategies for efficient estimation of soil organic content at the local scale based on a national spectral database. *Land Degradation & Development*, 33, pp. 1649-1661. DOI: 10.1002/ldr.4223
- Lehmann J., Kleber M. (2015). The contentious nature of soil organic matter. *Nature*, 528, pp. 60-68. DOI: 10.1038/nature16069
- Lobsey C.R., Viscarra Rossel R.A. (2016). Sensing of soil bulk density for more accurate carbon accounting. *European Journal of Soil Science*, 67, pp. 504-513. DOI: 10.1111/ejss.12355
- Ludwig B., Khanna P.K., Bauhus J., Hopmans P. (2002). Near infrared spectroscopy of forest soils to determine chemical and biological properties related to soil sustainability. *Forest Ecology and Management*, 171, pp. 121-132. DOI: 10.1016/S0378-1127(02)00467-X
- Ludwig B., Nitschke R., Terhoeven-Urselmans T., Michel K., Flessa H. (2008). Use of mid-infrared spectroscopy in the diffuse-reflectance mode for the prediction of the composition of organic matter in soil and litter. *Journal of Plant Nutrition and Soil Science*, 171, pp. 384-391. DOI: 10.1002/jpln.200700022
- Madari B.E., Reeves J.B. III, Coelho M.R., Machado P.L.O.A., De-Polli H., Coelho R.M., Benites V.M., Souza L.F., McCarty G.W. (2005). Mid- and near-infrared spectroscopic determination of carbon in a diverse set of soils from the Brazilian National Soil Collection. *Spectroscopy Letters*, 38, pp. 721-740. DOI: 10.1080/00387010500315876

- Madari B.E., Reeves J.B. III, Machado P.L.O.A., Guimaraes C.M., Torres E., McCarty G.W. (2006). Mid- and near-infrared spectroscopic assessment of soil compositional parameters and structural indices in two Ferralsols. *Geoderma*, 136, pp. 245-259. DOI: 10.1016/j.geoderma.2006.03.026
- Martin M.P., Arrouays D., Barré P., Boulonne L., Cécillon L., Chen S., Chenu C., Eglin T., Guenet B., Meersmans J., Mulder V.L., Saby N.P.A. (2022). Le Réseau de Mesures de la Qualité des Sols pour l'étude du carbone organique des sols en France métropolitaine. *Avancées scientifiques et applications. Le RMQS pour l'étude du carbone organique des sols. Étude et Gestion des Sols*, 29, pp. 239-351. https://www.afes.fr/wp-content/uploads/2022/07/EGS_2022_29_Martin_329-352.pdf (consulté le 24/10/2022).
- McCarty G.W., Reeves J.B. III, Reeves V.B., Follett R.F., Kimble J. (2002). Mid-infrared and near-infrared diffuse reflectance spectroscopy for soil carbon measurement. *Soil Science Society of America Journal*, 66, pp. 640-646. DOI: 10.2136/sssaj2002.0640
- McCarty G.W., Reeves J.B. III (2006). Comparison of near infrared and mid infrared diffuse reflectance spectroscopy for field-scale measurement of soil fertility parameters. *Soil Science*, 171, pp. 94-102. DOI: 10.1097/01.ss.0000187377.84391.54
- Minasny B., McBratney A.B., Tranter G., Murphy B.W. (2008). Using soil knowledge for the evaluation of mid-infrared diffuse reflectance spectroscopy for predicting soil physical and mechanical properties. *European Journal of Soil Science*, 59, pp. 960-971. DOI: 10.1111/j.1365-2389.2008.01058.x
- Moreira C.S., Brunet D., Verneyre L., Sá S.M.O., Galdos M.V., Cerri C.C., Bernoux M. (2009). Near infrared spectroscopy for soil bulk density assessment. *European Journal of Soil Science*, 60, pp. 785-791. DOI: 10.1111/j.1365-2389.2009.01170.x
- Morgan C.L.S., Waiser T.H., Brown D.J., Hallmark C.T. (2009). Simulated in situ characterization of soil organic and inorganic carbon with visible near-infrared diffuse reflectance spectroscopy. *Geoderma*, 151, pp. 249-256. DOI: 10.1016/j.geoderma.2009.04.010
- Mouazen A.M., De Baerdemaeker J., Ramon H. (2006). Effect of wavelength range on the measurement of some selected soil properties using visible-near infrared spectroscopy. *Journal of Near Infrared Spectroscopy*, 14, pp. 189-199. DOI: 10.1255/jnirs.614
- Nawar S., Mouazen A.M. (2018). Optimal sample selection for measurement of soil organic carbon using online vis-NIR spectroscopy. *Computers and Electronics in Agriculture*, 151, pp. 469-477. DOI: 10.1016/j.compag.2018.06.042
- Nawar S., Mouazen A.M. (2019). On-line vis-NIR spectroscopy prediction of soil organic carbon using machine learning. *Soil and Tillage Research*, 190, pp. 120-127. DOI: 10.1016/j.still.2019.03.006

- Ng W., Minasny B., Jones E., McBratney A. (2021). To spike or to localize? Strategies to improve the prediction of local soil properties using regional spectral library. *Geoderma*, 406, 115501. DOI: 10.1016/j.geoderma.2021.115501
- Nocita M., Kooistra L., Bachmann M., Müller A., Powell M., Weel S. (2011). Predictions of soil surface and topsoil organic carbon content through the use of laboratory and field spectroscopy in the Albany Thicket Biome of Eastern Cape Province of South Africa. *Geoderma*, 167-168, pp. 295-302. DOI: 10.1016/j.geoderma.2011.09.018
- Nocita M., Stevens A., Toth G., Panagos P., van Wesemael B., Montanarella L. (2014). Prediction of soil organic carbon content by diffuse reflectance spectroscopy using a local partial least square regression approach. *Soil Biology and Biochemistry*, 68, pp. 337-347. DOI: j.soilbio.2013.10.022
- O'Rourke S.M., Holden N.M. (2011). Optical sensing and chemometric analysis of soil organic carbon – a cost effective alternative to conventional laboratory methods? *Soil Use and Management*, 27, pp. 143-157. DOI: 10.1111/j.1475-2743.2011.00337.x
- Pansu M., Gautheyrou J., Loyer J.Y. (2001). *Soil analysis – Sampling, instrumentation, quality control*. Balkema Publishers, Lisse, Pays-Bas.
- Pansu M., Gautheyrou J. (2006). *Handbook of soil analysis – Mineralogical, organic and inorganic methods*. Springer, Heidelberg, Allemagne.
- Pei X., Sudduth K.A., Veum K.S., Li M. (2019). Improving in-situ estimation of soil profile properties using a multi-sensor probe. *Sensors*, 19, 1011. DOI: 10.3390/s19051011
- Pérez-Fernández E., Robertson A.H.J. (2016). Global and local calibrations to predict chemical and physical properties of a national spatial dataset of Scottish soils from their near infrared spectra. *Journal of Near Infrared Spectroscopy*, 24, pp. 293-316. DOI: 10.1255/jnirs.1229
- Poeplau C., Vos C., Don A. (2017). Soil organic carbon stocks are systematically overestimated by misuse of the parameters bulk density and rock fragment content. *SOIL*, 3, pp. 61-66. DOI: 10.5194/soil-3-61-2017
- Priori S., Fantappie M., Bianconi N., Ferrigno G., Pellegrini S., Costantini E.A.C. (2016). Field-scale mapping of soil carbon stock with limited sampling by coupling gamma-ray and Vis-NIR spectroscopy. *Soil Science Society of America Journal*, 80, pp. 954-964. DOI: 10.2136/sssaj2016.01.0018
- Rabenarivo M., Chapuis-Lardy L., Brunet D., Chotte J.L., Rabeharisoa L., Barthès B.G. (2013). Comparing near and mid-infrared reflectance spectroscopy for determining properties of Malagasy soils, using global or LOCAL calibration. *Journal of Near Infrared Spectroscopy*, 21, pp. 495-509. DOI: 10.1255/jnirs.1080
- Ramifehiarivo N., Barthès B.G., Cambou A., Chapuis-Lardy L., Chevallier T., Albrecht A., Razafimbelo T. (2023). Comparison of near and mid-infrared reflectance spectroscopy for the estimation of soil

- organic carbon fractions in Madagascar agricultural soils. *Accepted par Geoderma Regional*. DOI: 10.1016/j.geodrs.2023.e00638
- Ramirez-Lopez L., Behrens T., Schmidt K., Stevens A., Demattê J.A.M., Scholten T. (2013). The spectrum-based learner: A new local approach for modeling soil Vis–NIR spectra of complex datasets. *Geoderma*, 195-196, pp. 268-279. DOI: 10.1016/j.geoderma.2012.12.014
- Reeves J.B. III (2010). Near- versus mid-infrared diffuse reflectance spectroscopy for soil analysis emphasizing carbon and laboratory versus on-site analysis: Where are we and what needs to be done? *Geoderma*, 158, pp. 3-14. DOI: 10.1016/j.geoderma.2009.04.005
- Reeves J.B. III, McCarty G.W., Reeves V.B. (2001). Mid-infrared diffuse reflectance spectroscopy for the quantitative analysis of agricultural soils. *Journal of Agricultural and Food Chemistry*, 49, 766–772. DOI: 10.1021/jf0011283
- Reeves J.B. III, McCarty G.W., Mimmo T. (2002). The potential of diffuse reflectance spectroscopy for the determination of carbon inventories in soils. *Environmental Pollution*, 116, pp. S277-S284. DOI: 10.1016/S0269-7491(01)00259-7
- Roudier P., Hedley C.B., Ross C.W. (2015). Prediction of volumetric soil organic carbon from field-moist intact soil cores. *European Journal of Soil Science*, 66, pp. 651-660. DOI: 10.1111/ejss.12259
- Russell C.A. (2003). Sample preparation and prediction of soil organic matter properties by near infrared reflectance spectroscopy. *Communications in Soil Science and Plant Analysis*, 34, pp. 1557-1572. DOI: 10.1081/CSS-120021297
- Shen Z.F., Ramirez-Lopez L., Behrens T., Cui L., Zhang M.X., Walden L., Wetterlind J., Shi Z., Sudduth K.A., Song Y.Z., Catambay K., Viscarra Rossel R.A. (2022). Deep transfer learning of global spectra for local soil carbon monitoring. *ISPRS Journal of Photogrammetry and Remote Sensing*, 188, pp. 190-200. DOI: 10.1016/j.isprsjprs.2022.04.009
- Shepherd K.D., Walsh M.G. (2002). Development of reflectance spectral libraries for characterization of soil properties. *Soil Science Society of America Journal*, 66, pp. 988-998. DOI: 10.2136/sssaj2002.0988
- Shi Z., Ji R., Viscarra Rossel R.A., Chen S., Zhou Y. (2015). Prediction of soil organic matter using a spatially constrained local partial least squares regression and the Chinese vis–NIR spectral library. *European Journal of Soil Science*, 66, pp. 679-687. DOI: 10.1111/ejss.12272
- Siesler H.W. (2007). Basic principles of near-infrared spectroscopy. In Burns D.A., Ciurczak E.W. (Eds.), *Handbook of near-infrared analysis*, 3e édition, pp. 7-19. CRC Press, Boca Raton, Floride, USA.
- Socrates G. (2001). *Infrared and Raman characteristics group frequencies: Tables and charts*, 3e édition. John Wiley & Sons, Chichester, Royaume-Uni.

- Stenberg B. (2010). Effects of soil sample pretreatments and standardised rewetting as interacted with sand classes on Vis-NIR predictions of clay and soil organic carbon. *Geoderma*, 158, pp. 15-22. DOI: 10.1016/j.geoderma.2010.04.008
- Stenberg B., Viscarra Rossel R.A., Mouazen A.M., Wetterlind J. (2010). Visible and near infrared spectroscopy in soil science. *Advances in Agronomy*, 107, pp. 163-215. DOI: 10.1016/S0065-2113(10)07005-7
- Stevens A., van Wesemael B., Bartholomeus H., Rosillon D., Tychon B., Ben-Dor E. (2008). Laboratory, field and airborne spectroscopy for monitoring organic carbon content in agricultural soils. *Geoderma*, 144, pp. 395-404. DOI: 10.1016/j.geoderma.2007.12.009
- Stevens A., Nocita M., Tóth G., Montanarella L., van Wesemael B. (2013). Prediction of soil organic carbon at the European scale by visible and near infrared reflectance spectroscopy. *PLoS One*, 8, e66409. DOI: 10.1371/journal.pone.0066409
- Sudduth K.A., Hummel J.W. (1991). Evaluation of reflectance methods for soil organic matter sensing. *Transactions of the ASAE*, 34, pp. 1900-1909.
- Sudduth K.A., Hummel J.W. (1993). Soil organic-matter, CEC, and moisture sensing with a portable NIR spectrophotometer. *Transactions of the ASAE*, 36, pp. 1571-1582.
- Terhoeven-Urselmans T., Michel K., Helfrich M., Flessa H., Ludwig B. (2006). Near-infrared spectroscopy can predict the composition of organic matter in soil and litter. *Journal of Plant Nutrition and Soil Science*, 169, pp. 168-174. DOI: 10.1002/jpln.200521712
- Tóth G., Jones A., Montanarella L. (2013). The LUCAS topsoil database and derived information on the regional variability of cropland topsoil properties in the European Union. *Environmental Monitoring and Assessment*, 185, pp. 7409-7425. DOI: 10.1007/s10661-013-3109-3
- Twomey S.A., Bohren C.F., Mergenthaler J.L. (1986). Reflectance and albedo differences between wet and dry surfaces. *Applied Optics*, 25, pp. 431-437. DOI: 10.1364/AO.25.000431
- Vågen T.G., Winowiecki L.A., Tondoh J.E., Desta L.T., Gumbrecht, T. (2016). Mapping of soil properties and land degradation risk in Africa using MODIS reflectance. *Geoderma*, 263, pp. 216-225. DOI: 10.1016/j.geoderma.2015.06.023
- Van Groenigen J.W., Muters C.S., Horwath W.R., van Kessel C. (2003). NIR and DRIFT-MIR spectrometry of soils for predicting soil and crop parameters in a flooded field. *Plant and Soil*, 250, pp. 155-165. DOI: 10.1023/A:1022893520315
- Veum K.S., Sudduth K.A., Kremer R.J., Kitchen N.R. (2015). Estimating soil quality index with VNIR reflectance spectroscopy. *Soil Science Society of America Journal*, 79, pp. 637-649. DOI: 10.2136/sssaj2014.09.0390

- Veum K.S., Parker P.A., Sudduth K.A., Holan S.H. (2018). Predicting profile soil properties with reflectance spectra via Bayesian covariate-assisted external parameter orthogonalization. *Sensors*, 18, 3869. DOI: 10.3390/s18113869
- Viscarra Rossel R.A., Walvoort D.J.J., McBratney A.B., Janik L.J., Skjemstad J.O. (2006). Visible, near-infrared, mid-infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, 131, pp. 59-75. DOI: 10.1016/j.geoderma.2005.03.007
- Viscarra Rossel R.A., Adamchuk V.I., Sudduth K.A., McKenzie N.J., Lobsey C. (2011). Proximal soil sensing. An effective approach for soil measurements in space and time. *Advances in Agronomy*, 113, pp. 237-282. DOI: 10.1016/B978-0-12-386473-4.00010-5
- Viscarra Rossel R.A., Webster R. (2012). Predicting soil properties from the Australian soil visible–near infrared spectroscopic database. *European Journal of Soil Science*, 63, pp. 848-860. DOI: 10.1111/j.1365-2389.2012.01495.x
- Viscarra Rossel R.A., Behrens T., Ben-Dor E., Brown D.J., Demattê J.A.M., Shepherd K.D., Shi Z., Stenberg B., Stevens A., Adamchuk V., Aichi H., Barthès B.G., Bartholomeus H.M., Bayer A.D., Bernoux M., Böttcher K., Brodský L., Changwen D., Chappell A., Fouad Y., Genot V., Gomez C., Grunwald S., Gubler A., Guerrero C., Hedley C.B., Knadel M., Morrás H.J.M., Nocita M., Ramirez-Lopez L., Roudier P., Rufasto Campos E.M., Sanborn P., Sellitto V.M., Sudduth K.A., Rawlins B.G., Walter C., Winowiecki L.A., Young Hong S., Ji W. (2016a). A global spectral library to characterize the world's soil. *Earth-Science Reviews*, 155, pp. 198–230. DOI: 10.1016/j.earscirev.2016.01.012
- Viscarra Rossel R.A., Brus D.J., Lobsey C., Shi Z., McLachlan G. (2016b). Baseline estimates of soil organic carbon by proximal sensing: Comparing design-based, model-assisted and model-based inference. *Geoderma*, 265, pp. 152-163. DOI: 10.1016/j.geoderma.2015.11.016
- Viscarra Rossel R.A., Lobsey C.R., Sharman C., Flick P., McLachlan G. (2017). Novel environmental sensing for monitoring soil organic C stocks and condition. *Environmental Science & Technology*, 51, pp. 5630-5641. DOI: 10.1021/acs.est.7b00889
- Wijewardane N.K., Ge Y.F., Sanderman J., Ferguson R. (2021). Fine grinding is needed to maintain the high accuracy of mid-infrared diffuse reflectance spectroscopy for soil property estimation. *Soil Science Society of America Journal*, 85, pp. 263-272. DOI: 10.1002/saj2.20194
- Workman J. Jr, Weyer L. (2008). *Practical guide to interpretative near-infrared spectroscopy*. CRC Press, Boca Raton, Florida, USA.
- Xie H.T., Yang X.M., Drury C.F., Yang J.Y., Zhang X.D. (2011). Predicting soil organic carbon and total nitrogen using mid- and near-infrared spectra for Brookston clay loam soil in Southwestern Ontario. *Canadian Journal of Soil Science*, 91, pp. 53-63. DOI: 10.4141/CJSS10029
- Yang X.M., Xie H.T., Drury C.F., Reynolds W.D., Yang J.Y., Zhang X.D. (2012). Determination of organic carbon and nitrogen in particulate organic matter and particle size fractions of Brookston clay loam

soil using infrared spectroscopy. *European Journal of Soil Science*, 63, pp. 177-188. DOI: 10.1111/j.1365-2389.2011.01421.x

Zimmermann M., Leifeld J., Fuhrer J. (2007). Quantifying soil organic carbon fractions by infrared-spectroscopy. *Soil Biology and Biochemistry*, 39, pp. 224-231. DOI:10.1016/j.soilbio.2006.07.010

Tableau supplémentaire 1 : Prédications de la concentration en carbone organique du sol à partir de grosses bibliothèques spectrales (> 1000 échantillons ; spectres d'échantillons séchés à l'air < 2 mm, sauf Gogé *et al.*, 2012 et 2014, < 1 mm, et Grinand *et al.*, 2012, et Clairotte *et al.*, 2016, < 0,2 mm).

Supplementary table 1: Predictions of soil organic carbon concentration from large spectral libraries (> 1000 samples; spectra of air-dried samples < 2 mm, except Gogé *et al.*, 2012 and 2014, < 1 mm, and Grinand *et al.*, 2012, and Clairotte *et al.*, 2016, < 0.2 mm).

Référence	Gamme	Librairie	Echantillons de validation	Approche	R ² val _i	RPDval
Shepherd et Walsh (2002)	VPIR	> 1000 éch. 0-20 cm, 7 pays est-africains	33 % au hasard	MARS	0,80	2,2
Brown <i>et al.</i> (2006)	VPIR	> 4100 éch. plusieurs profondeurs, surtout USA	Valid. croisée, 6 groupes	BRT	0,82/0,87*	2,4/2,8*
Viscarra Rossel et Webster (2012)	VPIR	> 10000 éch., plusieurs profondeurs, Australie	33 % les mieux représentés	RT	0,79**	2,2**
Gogé <i>et al.</i> (2012)	VPIR	> 2100 éch. 0-30 cm, toute la France	33 % les mieux représentés	PLS locale pondérée	0,86	2,7
Grinand <i>et al.</i> (2012)	MIR	> 2100 éch. 0-30 cm, toute la France	80 % au hasard, 10 répét.	PLS	0,89	3,0
Gogé <i>et al.</i> (2014)	VPIR	> 2100 éch. 0-30 cm, toute la France	Indépendants, 24 km ²	PLS locale pond.+ spik.	0,58	1,4
Shi <i>et al.</i> (2015)	VPIR	> 2700 éch. 0-20 cm, 20 provinces chinoises	10 % bien représentés	PLS	0,50	1,4
	VPIR	> 2700 éch. 0-20 cm, 20 provinces chinoises	10 % bien représentés	PLS locale	0,69	1,8
	VPIR	> 2700 éch. 0-20 cm, 20 provinces chinoises	10 % bien représentés	PLS locale x 2	0,74	2,0
Clairotte <i>et al.</i> (2016)	PIR	> 3800 éch. 0-30 et 30-50 cm, toute la France	10 % les mieux représentés	PLS	0,85	2,0
	PIR	> 3800 éch. 0-30 et 30-50 cm, toute la France	10 % les mieux représentés	PLS locale	0,89	3,0
	MIR	> 3800 éch. 0-30 et 30-50 cm, toute la France	10 % les mieux représentés	PLS	0,88	2,5
	MIR	> 3800 éch. 0-30 et 30-50 cm, toute la France	10 % les mieux représentés	PLS locale	0,92	3,4
Stevens <i>et al.</i> (2013)		> 19000 éch. 0-30 cm, 23 pays européens	33 % les mieux représentés			
	VPIR	Cultures		SVR	0,67/0,79*	1,7/2,2*
	VPIR	Prairies		SVR	0,71/0,87*	1,9/2,7*
	VPIR	Forêts		SVR	0,75/0,89*	2,0/2,9*
	VPIR	Autres sols minéraux		SVR	0,78/0,86*	2,1/2,6*
	VPIR	Sols organiques		Cubist	0,76/0,76*	2,0/2,0*
Nocita <i>et al.</i> (2014)		> 19000 éch. 0-30 cm, 23 pays européens	30 % les mieux représentés			
	VPIR	Cultures		PLS locale	0,79/0,84*	2,1/2,5*
	VPIR	Prairies		PLS locale	0,81/0,84*	2,3/2,5*
	VPIR	Forêts		PLS locale	0,79/0,85*	2,1/2,5*
	VPIR	Sols organiques		PLS locale	0,76/0,76*	2,0/2,0*
Viscarra Rossel <i>et al.</i> (2016a)	VPIR	> 23000 éch., plusieurs profondeurs, 5 continents	25 % au hasard	Cubist après classif.	0,89	3,3

* Teneur en sables utilisée comme covariable (pour la seconde valeur). ** Prédiction pour log₁₀COS. PLS, régression des moindres carrés partiels ; PLS locale, utilisant des étalons voisins spectraux ; PLS locale x 2, utilisant des voisins spectraux et géographiques ; PLS locale pondérée ou pond. (la contribution des voisins spectraux est pondérée selon leur proximité avec l'échantillon cible) ; spik., spiking, enrichissement de la librairie avec quelques échantillons du site cible ; MARS, régression multivariée par spline adaptative ; BRT, arbre de régression renforcé ; Cubist, arbre de décision et régression ; SVR, régression à vecteurs de support.

Tableau supplémentaire 2 : Prédictions de la concentration en COS à partir de spectres VPIR ou MIR acquis in situ, et (italiques) éventuelles prédictions homologues à partir de spectres de sol sec < 2 mm acquis avec le même instrument. Régression PLS sauf Nawar et Mouazen (2019 ; random forest).

Supplementary table 2: Predictions of SOC concentration using VNIR or MIR spectra acquired in the field, and (in italics) possible homologous predictions using spectra acquired on dried samples < 2 mm with the same instrument. PLS regression for all studies except Nawar and Mouazen (2019; random forest).

Référence	Gamme	Conditions d'acquisition des spectres	Nombre d'éch. et de niveaux de profondeur	Echelle	Type de validation	R ²	RPD
Mouazen <i>et al.</i> (2006)	VPIR <i>VPIR</i>	Sol frais émietté <i>Sol sec < 2 mm</i>	186, 1	Multi-site, régional	LOO CV	0,76 <i>0,80</i>	2,0 2,2
Stevens <i>et al.</i> (2008)	VPIR <i>VPIR</i>	A 1 m de hauteur <i>Sol sec < 2 mm</i>	117, 1	10 champs, régional	LOO CV	0,76 <i>0,75</i>	2,1 2,0
Kusumo <i>et al.</i> (2008)	VPIR	Carotte sol intact	210, 2	7 champs, régional	Externe hasard	0,75	2,0
Morgan <i>et al.</i> (2009)	VPIR <i>VPIR</i>	Carotte sol intact <i>Sol sec < 2 mm</i>	540, 7	6 champs, régional	Externe hasard	0,64 <i>0,73</i>	1,5 1,7
Kusumo <i>et al.</i> (2011)	VPIR	Carotte sol intact	90, 5	1 champ	LOO CV	0,86	2,7
Nocita <i>et al.</i> (2011)	VPIR <i>VPIR</i>	Surface du sol <i>Sol sec < 2 mm</i>	113, 1	transect de 130 km	Externe représ.	0,84 <i>0,93</i>	2,5 3,7
Izaurrealde <i>et al.</i> (2013)	MIR	Carotte émiettée	110+spiking, 3	1 champ	Externe	0,77	2,1
Gras <i>et al.</i> (2014)	VPIR <i>VPIR</i> VPIR <i>VPIR</i>	Carotte sol perturbé Surface du sol Carotte émiettée <i>Sol sec < 2 mm</i>	201, 1 201, 1 201, 1	6 champs en France 6 champs en France 6 champs en France	CV 4 groupes CV 4 groupes CV 4 groupes	0,86 0,82 0,83 <i>0,80</i>	2,8 2,4 2,4 2,3
Cambou <i>et al.</i> (2016)	VPIR	Carotte sol perturbé	288, 3	2 champs, 250 km entre	Externe représ.	0,75	2,0
Ji <i>et al.</i> (2016)	MIR <i>MIR</i> MIR <i>MIR</i>	A 5 cm de prof. <i>Sol sec < 2 mm</i> A 5 cm de prof. <i>Sol sec < 2 mm</i>	120, 1 121, 1	1 champ 1 autre champ	LOO CV LOO CV	0,86 <i>0,90</i> 0,60 <i>0,61</i>	2,6 3,2 1,6 1,6
Nawar et Mouazen (2018)	VPIR <i>VPIR</i>	A 15 cm de prof. <i>Sol sec < 2 mm</i>	268+spiking, 1	4 champs, régional	Externe représ.	0,78 <i>0,84</i>	2,1 2,5
Veum <i>et al.</i> (2018)	VPIR <i>VPIR</i>	Trou de sonde <i>Sol sec < 2 mm</i>	708, 5	22 champs, 2 Etats (US)	Externe hasard	0,64 <i>0,82</i>	1,6 2,3
Allory <i>et al.</i> (2019)	VPIR <i>VPIR</i>	Face de fosse <i>Sol sec < 2 mm</i>	132, 2 à 7	5 sites, 850 km entre	Externe représ.	0,83 <i>0,95</i>	2,4 4,6
Hutengs <i>et al.</i> (2019)	VPIR <i>VPIR</i> MIR <i>MIR</i>	Surface du sol <i>Sol sec < 2 mm</i> Surface du sol <i>Sol sec < 2 mm</i>	90, 1 90, 1	90 champs, régional 90 champs, régional	Externe représ. Externe représ.	0,27 <i>0,66</i> 0,63 <i>0,77</i>	1,2 1,3 1,6 2,0
Nawar et Mouazen (2019)	VPIR VPIR VPIR	A 15 cm de prof. A 15 cm de prof. A 15 cm de prof.	122, 1 139, 1 240+spiking, 1	1 champ 1 autre champ 4 champs, 300 km entre	Externe représ. Externe représ. Externe représ.	0,65 0,71 0,75	1,7 1,9 2,0
Pei <i>et al.</i> (2019)	VPIR VPIR VPIR	Trou de sonde Trou de sonde Trou de sonde	90, 4 58, 4 148, 4	1 champ 1 autre champ 2 champs, 3 km entre	Externe hasard Externe hasard Externe hasard	0,77 0,87 0,80	2,1 2,7 2,2
Allo <i>et al.</i> (2020)	VPIR <i>VPIR</i>	Carotte sol intact <i>Sol sec < 2 mm</i>	95, ≈ 10	11 sites, régional	LOO CV	0,93 <i>0,92</i>	3,7 3,6
Greenberg <i>et al.</i> (2021)	VPIR MIR	Surface du sol Surface du sol	238, 1	1 champ	Externe hasard	n.d.	2,3 1,9
Greenberg <i>et al.</i> (2022)	MIR	Surface du sol	450, 1	3 champs, multilocal	Externe indép. Externe représ.	n.d. n.d.	≈ 2 ≈ 3

CV, validation croisée ; LOO CV, *leave-one-out* CV (chaque échantillon prédit à partir des autres) ; Ext., validation externe avec ≈ 30 % des échantillons en général, sélectionnés au hasard, ou représentés par les étalons (cas noté « représ. » ; par exemple un échantillon sur trois après tri par COS croissant), ou validation indépendante (cas noté « indép. » ; par exemple site éloigné) ; spiking, enrichissement de la librairie avec quelques échantillons du site cible ; n.d. non déterminé.

